

# Anomalous diffusion by fractal homogenization

Scott Armstrong <sup>\*</sup> Vlad Vicol <sup>†</sup>

revised version of October 9, 2024

## Abstract

For every  $\alpha < 1/3$ , we construct an explicit divergence-free vector field  $\mathbf{b}(t, x)$  which is periodic in space and time and belongs to  $C_t^0 C_x^\alpha \cap C_t^\alpha C_x^0$  such that the corresponding scalar advection-diffusion equation

$$\partial_t \theta^\kappa + \mathbf{b} \cdot \nabla \theta^\kappa - \kappa \Delta \theta^\kappa = 0$$

exhibits anomalous dissipation of scalar variance for arbitrary  $H^1$  initial data:

$$\limsup_{\kappa \rightarrow 0} \int_0^1 \int_{\mathbb{T}^d} \kappa |\nabla \theta^\kappa(t, x)|^2 dx dt > 0.$$

The vector field is deterministic and has a fractal structure, with periodic shear flows alternating in time between different directions serving as the base fractal. These shear flows are repeatedly inserted at infinitely many scales in suitable Lagrangian coordinates. Using an argument based on ideas from quantitative homogenization, the corresponding advection-diffusion equation with small  $\kappa$  is progressively renormalized, one scale at a time, starting from the (very small) length scale determined by the molecular diffusivity up to the macroscopic (unit) scale. At each renormalization step, the effective diffusivity is enhanced by the influence of advection on that scale. By iterating this procedure across many scales, the effective diffusivity on the macroscopic scale is shown to be of order one.

## Contents

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>Introduction</b>  | <b>2</b>  |
| 1.1      | Motivation and prior results on anomalous dissipation of scalar variance | 4         |
| 1.2      | An outline of the proof: fractal homogenization                          | 9         |
| 1.3      | Uniform estimates for the scalar   | 16        |
| 1.4      | Notation   | 17        |
| 1.5      | Outline of the paper   | 17        |
| <b>2</b> | <b>The fractal vector field: construction and regularity</b>             | <b>18</b> |
| 2.1      | A list of the ingredients used in the construction                       | 18        |
| 2.2      | Construction of the vector field   | 22        |
| 2.3      | Regularity of the stream functions and associated flows                  | 24        |
| 2.4      | Material derivative estimates  | 30        |

<sup>\*</sup>Courant Institute of Mathematical Sciences, New York University. [scotta@cims.nyu.edu](mailto:scotta@cims.nyu.edu).

<sup>†</sup>Courant Institute of Mathematical Sciences, New York University. [vicol@cims.nyu.edu](mailto:vicol@cims.nyu.edu).

|                   |   |            |
|-------------------|---|------------|
| <b>3</b>          | <b>Correctors and renormalized diffusivities</b>                                    | <b>36</b>  |
| 3.1               | The correctors: definitions and estimates . . . . .                                 | 36         |
| 3.2               | Correctors and effective diffusivity for smoothly alternating shear flows . . . . . | 36         |
| 3.3               | The renormalized diffusivities: recurrence, averaging and estimates . . . . .       | 38         |
| <b>4</b>          | <b>The multiscale ansatz</b>  | <b>45</b>  |
| 4.1               | Ingredients for the multiscale ansatz . . . . .                                     | 48         |
| 4.2               | Definition of the multiscale ansatz $\tilde{\theta}_m$ . . . . .                    | 51         |
| 4.3               | Regularity estimates for $T_{m-1}$ . . . . .  | 52         |
| 4.4               | Estimates for $\tilde{H}_m$ . . . . .   | 65         |
| <b>5</b>          | <b>Homogenization cascade up the inertial-convection subrange</b>                   | <b>68</b>  |
| 5.1               | Computing the error in the multiscale ansatz . . . . .                              | 68         |
| 5.2               | Estimates of the nine error terms . . . . .   | 73         |
| 5.3               | Energy cascade down the scales . . . . .  | 82         |
| 5.4               | The proof of Theorem 1.1 . . . . .  | 90         |
| 5.5               | Lack of selection in the vanishing diffusivity limit . . . . .                      | 93         |
| <b>A</b>          | <b>Macroscopic mean drift destroys enhancement</b>                                  | <b>100</b> |
| <b>B</b>          | <b>Faá di Bruno formula and its consequences</b>                                    | <b>103</b> |
| B.1               | Useful lemmas . . . . .   | 103        |
| B.2               | Application to the transport equation . . . . .                                     | 106        |
| B.3               | Application to ODEs . . . . .   | 109        |
| <b>C</b>          | <b>Ergodic lemmas for periodic functions</b>  | <b>112</b> |
| C.1               | Some basic ergodic lemmas in one dimension . . . . .                                | 112        |
| <b>References</b> |   | <b>116</b> |

## 1. Introduction

We consider the Cauchy problem for the linear advection-diffusion equation

$$\begin{cases} \partial_t \theta^\kappa + \mathbf{b} \cdot \nabla \theta^\kappa - \kappa \Delta \theta^\kappa = 0 & \text{in } (0, \infty) \times \mathbb{T}^d, \\ \theta^\kappa(0, \cdot) = \theta_0 & \text{on } \mathbb{T}^d. \end{cases} \quad (1.1)$$

The initial data  $\theta_0$  is assumed to belong to  $L^2(\mathbb{T}^d)$  and have zero mean; it can also be assumed to be smooth. The vector field  $\mathbf{b}(t, x)$  in (1.1) is assumed to be *incompressible*, that is, divergence-free:

$$\nabla \cdot \mathbf{b}(t, \cdot) = 0, \quad \forall t \in (0, \infty). \quad (1.2)$$

Physically, the solution  $\theta^\kappa$  represents a scalar quantity, such as temperature or the concentration of a pollutant in a fluid, which is “passive” in the sense of having a negligible effect on the flow itself. For this reason, the equation in (1.1) is often called the *passive scalar equation*. We are interested in the case in which the parameter  $\kappa > 0$  is very small and the vector field  $\mathbf{b}(t, x)$ , although continuous in  $(t, x)$ , is still quite rough—possessing certain properties characteristic of turbulent flows.

The main result of this paper is the construction of an explicit vector field  $\mathbf{b}(t, x)$  for which the variance of the corresponding passive scalar  $\theta^\kappa$  exhibits anomalous dissipation.

**Theorem 1.1** (Anomalous dissipation of scalar variance). *Let  $d \geq 2$  and  $\alpha \in (0, 1/3)$ . There exists a vector field*

$$\mathbf{b} \in C_t^0 C_x^{0,\alpha}([0, 1] \times \mathbb{T}^d) \cap C_t^{0,\alpha} C_x^0([0, 1] \times \mathbb{T}^d) \quad (1.3)$$

*which satisfies (1.2) such that, for every mean-zero initial datum  $\theta_0 \in H^1(\mathbb{T}^d)$ , the family of unique solutions  $\{\theta^\kappa\}_{\kappa>0} \in C([0, 1]; L^2(\mathbb{T}^d))$  of the advection-diffusion equation (1.1) satisfy*

$$\limsup_{\kappa \rightarrow 0} \kappa \|\nabla \theta^\kappa\|_{L^2((0,1) \times \mathbb{T}^d)}^2 \geq \varrho^2 \|\theta_0\|_{L^2(\mathbb{T}^d)}^2, \quad (1.4)$$

*for some constant  $\varrho = \varrho(d, \theta_0) \in (0, 1]$  which depends only on  $d$  and the ratio  $\|\theta_0\|_{L^2(\mathbb{T}^d)} / \|\nabla \theta_0\|_{L^2(\mathbb{T}^d)}$ .*

The initial-value problem (1.1) has a unique global solution for every  $\kappa > 0$  provided that the vector field  $\mathbf{b}(t, x)$  belongs to  $L_t^\infty L_x^d$ . By the incompressibility condition (1.2), this solution satisfies the energy balance relation

$$\|\theta_0(\cdot)\|_{L^2(\mathbb{T}^d)}^2 - \|\theta^\kappa(1, \cdot)\|_{L^2(\mathbb{T}^d)}^2 = 2\kappa \|\nabla \theta^\kappa\|_{L^2((0,1) \times \mathbb{T}^d)}^2. \quad (1.5)$$

The quantity on the right side of (1.5) is therefore called the *dissipation of scalar variance*. While norms of  $\mathbf{b}(t, x)$  do not appear explicitly in (1.5), the solution  $\theta^\kappa$  of course depends on the vector field in a very complicated and nonlinear way.

The family  $\{\theta^\kappa\}_{\kappa>0}$  in Theorem 1.1 are actually classical solutions of (1.1). Indeed, the incompressibility condition (1.2) allows us to write the drift term as part of the second-order diffusion term, using a stream matrix which, in view of (1.3), belongs to  $C^{1,\alpha}$ . The standard Schauder estimates therefore imply that, for each  $\kappa > 0$ , the solution of (1.1) belongs to  $C_t^0 C_x^{2,\alpha} \cap C_t^{1,\alpha/2} C_x^0$  for positive times.

As we will see in the proof, the parameter  $\varrho > 0$  in Theorem 1.1 can be taken to be

$$\varrho(d, \theta_0) = c \left( \frac{\|\theta_0\|_{L^2(\mathbb{T}^d)}}{\|\theta_0\|_{H^1(\mathbb{T}^d)}} \right)^{\frac{1+\alpha}{1-\alpha} + \varepsilon}, \quad (1.6)$$

where  $\varepsilon > 0$  is any positive constant, and  $c = c(d, \varepsilon) > 0$  is a positive constant. In particular,  $\varrho(d, \theta_0)$  depends only on  $d$  and a lower bound for the length scale  $\|\theta_0\|_{L^2(\mathbb{T}^d)} / \|\nabla \theta_0\|_{L^2(\mathbb{T}^d)}$ .

The proof of Theorem 1.1 gives more information concerning the scalar  $\theta^{\kappa_j}$  along the subsequence  $\kappa_j \downarrow 0$  than what appears in the statement. For instance, it yields a tiny exponent  $\mu > 0$  depending only on  $\alpha$  such that  $\|\theta^{\kappa_j}\|_{C_t^{0,\mu} L_x^2([0,1] \times \mathbb{T}^d)}$  stays bounded along the subsequence  $\kappa_j \downarrow 0$ . The mapping  $t \mapsto \int_{\mathbb{T}^d} |\theta^{\kappa_j}(t, x)|^2 dx$  is therefore uniformly continuous on  $[0, 1]$ : see Remark 5.4. In particular, the diffusive anomaly does not happen at any single “blow-up time.” The uniform regularity of the scalar  $\theta^{\kappa_j}$  will be greatly improved in a forthcoming paper [ARV].

Our arguments exhibit an explicit subsequence  $\kappa_j \rightarrow 0$  along which the lower bound in (1.4) is realized, which, in particular, does not depend on  $\theta_0$ . In fact, we construct a sequence of disjoint intervals  $I_j := [\frac{1}{2}\kappa_j, 2\kappa_j]$  with  $\kappa_j \rightarrow 0$  such that, if  $I = \cup_{j \in \mathbb{N}} I_j$ , then

$$\inf_{\kappa \in I} \kappa \|\nabla \theta^\kappa\|_{L^2((0,1) \times \mathbb{T}^d)}^2 \geq \varrho^2 \|\theta_0\|_{L^2(\mathbb{T}^d)}^2.$$

The position of  $\kappa$  within the interval  $I_j$  determines, up to an error which can be made arbitrarily small, the value of  $\kappa \|\nabla \theta^\kappa\|_{L^2((0,1) \times \mathbb{T}^d)}^2$ , with the left and right endpoints of  $I_j$  giving rise to significantly different values, at least for carefully chosen initial data. We therefore demonstrate the lack of a selection principle for vanishing diffusivity limits to the solutions of the transport equation

( $\kappa = 0$  in (1.1)), which evidently possesses non-unique bounded weak solutions. For a precise statement see Proposition 5.5, which is established in Section 5.5.

The vector field  $\mathbf{b}(t, x)$  appearing in the statement of Theorem 1.1 has an explicit construction using deterministic ingredients, namely periodic shear flows with directions that alternate in time. An infinite sequence of copies of these shear flows are embedded in the vector field, each with a different wave number, with the sequence of wave numbers tending to infinity at a super-geometric rate. The proof of Theorem 1.1 is based on a renormalization of effective diffusivities, in which each active scale in the vector field is homogenized, one-by-one. Each homogenization step enhances the effective diffusivity of the equation. After an iteration up the scales, this reveals an effective diffusivity of order one on the macroscopic scale, which implies anomalous diffusion. In Section 1.2, below, we review the motivation for the construction of the vector field and give an outline of the reiterated homogenization method used to prove Theorem 1.1. In Section 1.3 we discuss future extensions of this result.

Theorem 1.1 and its proof provide only an example of a vector field  $\mathbf{b}(t, x)$  such that the advection-diffusion equation displays anomalous diffusion. Examples—the more physically realistic the better—are certainly useful for building intuition about very complex phenomena. However, we believe that the main value of this work is in the proof strategy, which is a demonstration of the possibility of rigorously proving anomalous diffusion by analyzing the backwards cascade of eddy diffusivities via quantitative homogenization techniques. We expect this point of view to be robust and of independent interest to the broader area of rigorous hydrodynamic turbulence.

## 1.1 Motivation and prior results on anomalous dissipation of scalar variance

If the vector field  $\mathbf{b}(t, x)$  has significantly more spatial regularity than (1.3)—for example, if it belongs to  $L_t^1 C_x^{0,1}$ —then the flows determined by the vector field are well-defined and the corresponding transport equation is well-posed (by standard ODE theory), which then must be the equation satisfied by the limit as  $\kappa \rightarrow 0$  of the solutions  $\theta^\kappa$ . Consequently, as the flows are measure-preserving by (1.2), we deduce that

$$\lim_{\kappa \rightarrow 0} \|\theta^\kappa(t, \cdot)\|_{L^2(\mathbb{T}^d)} = \|\theta_0\|_{L^2(\mathbb{T}^d)}, \quad \forall t \in (0, \infty). \quad (1.7)$$

In view of (1.5), this limit is equivalent to

$$\lim_{\kappa \rightarrow 0} \kappa \|\nabla \theta^\kappa\|_{L^2((0,t) \times \mathbb{T}^d)}^2 = 0, \quad \forall t \in (0, \infty), \quad (1.8)$$

which is evidently in contrast to the conclusion of Theorem 1.1.

If the limit in (1.8) does not hold, then we speak of *anomalous dissipation of scalar variance* or, alternatively, *anomalous diffusion*. It is widely expected that solutions of (1.1) with vector fields  $\mathbf{b}(t, x)$  which are rougher than Lipschitz in space (but still Hölder continuous) may exhibit anomalous dissipation of scalar variance. This prediction was first discussed by Obukhov in [Obu49].

Indeed, anomalous diffusion is presumed to occur for vector fields describing the velocity of a turbulent fluid, and is a basic assumption in phenomenological theories of scalar turbulence in the physics literature. This remarkable prediction that the rate of dissipation is independent of  $\kappa$ , when  $\mathbf{b}(t, x)$  describes a turbulent flow, is backed by very strong experimental and numerical evidence [SS00, War00, DSY05].

The reason that anomalous diffusion is expected to hold for “turbulent” velocity fields is explained in the physics literature roughly as follows. A characteristic of a turbulent velocity

field  $\mathbf{b}(t, x)$  is that it exhibits activity across a large range of length scales. Advection by the velocity field rearranges the level sets of the scalar  $\theta^\kappa$ , creating wiggles on smaller length scales, which are then mixed by the features of the velocity field on those smaller scales. This process continues across a large number of length scales, called the *inertial-convection range*, with smaller and smaller spatial oscillations created. Finally, the wiggles in the scalar reach down to the very small scale at which the molecular diffusivity dominates advection, at which point they are dissipated away.

A rigorous theoretical explanation of this phenomenon is still elusive. In fact, the mathematical analysis seems to lag the phenomenological theories by so much that not even a *satisfactory example* of anomalous dissipation for passive scalars is available (the few available results are discussed in Section 1.1, below).

It is not hard to see why this is so: the physicists' explanation is the only way anomalous dissipation can happen. For very small  $\kappa$ , the diffusion term  $\kappa\Delta$  essentially acts only on very small length scales—otherwise its effect is negligible and the advection term dominates. But it is clear from the identity (1.5) that the diffusion term is the only thing can be responsible for dissipation. If anomalous dissipation is observed, it must be the vector field that is responsible for pushing the oscillations of the scalar into smaller and smaller scales. Since wiggles in the vector field interact with those of the scalar only if their wave numbers are separated by at most an order of magnitude,<sup>1</sup> it follows that both the vector field and the scalar must have a large number of active scales whose interactions span the range from the macroscopic scale to the “inertial” scale on which the diffusion is felt. Since the  $\theta^\kappa$  depends on  $\mathbf{b}(t, x)$  in a highly nontrivial, nonlinear fashion, it is very challenging to analyze such a situation—even if one is permitted to construct the vector field.

There are essentially only two known classes of examples which exhibit anomalous dissipation of scalar variance. The first is a stochastic model which is very rough in time (the Kraichnan model), and the second is a class of deterministic vector fields which are “quasi self-similar” and have only one active scale at each time (the singularly focusing alternating shear flows).

**The Kraichnan model.** Kraichnan introduced in [Kra68] a simplified model for passive scalar turbulence, one of the early examples of “synthetic turbulence”. He proposed that  $\mathbf{b} = \mathbf{b}^\nu$  is taken to be a realization of a statistically homogeneous, isotropic, stationary Gaussian random field, which has zero mean, is very rough in time (it has white-noise correlation), and is colored in space (with a Kolmogorov-type scaling of increments in space, above a certain scale).<sup>2</sup> Then one is to study the statistics of the field  $\theta^\kappa$  solving (1.1) (understood in the Stratonovich sense,  $d\theta^\kappa - \kappa\Delta\theta^\kappa dt = d\mathbf{b}^\nu \circ \nabla\theta^\kappa$ ). The main result concerning anomalous diffusion (1.4) in the joint  $\nu, \kappa \rightarrow 0$  limit, was established by Bernard, Gawedzki, and Kupiainen [BGK98]. See also [GV00, VEE00, EVE01, LJR02] for further results and refinements. Moreover, the Lagrangian flows  $\xi^{\nu, \kappa}$  become non-unique and stochastic in the  $\nu, \kappa \rightarrow 0$  limit, for a fixed initial particle position and a fixed velocity realization  $\mathbf{b}^\nu$ . This phenomenon is called *spontaneous stochasticity*. In fact, it was shown by Drivas and Eyink [DE17] that spontaneous stochasticity is equivalent to anomalous dissipation, not just for the Kraichnan model, but for any passive scalar transport of the type (1.1)

<sup>1</sup>This is due to the incompressibility condition (1.2), and the implicit assumption that  $\mathbf{b}(t, x)$  is continuous. If the constraint (1.2) is dropped, then it is easy to make examples, for instance by creating a vector field which pushes all particles into a small neighborhood of the origin before suddenly pushing them away in radial directions. If the vector field is allowed to be very rough in time, then small scales can also be created fairly easily, as discussed below.

<sup>2</sup>More precisely,  $\mathbf{b} = \mathbf{b}^\nu$  (here  $\nu$  denotes an inverse Reynolds number) has covariance  $\langle(\mathbf{b}_i^\nu(x, t) - \mathbf{b}_i^\nu(y, t))(\mathbf{b}_j^\nu(x, s) - \mathbf{b}_j^\nu(y, s))\rangle = \delta(t - s)D_{ij}(x - y)$ , where the matrix  $D_{ij}$  is symmetric, it has incompressible rows  $\partial_j D_{ij} = 0$ , and most importantly, the diagonal entries satisfy  $D_{ii}(z) = A|z|^{2\alpha}$  for  $\ell_\nu \ll |z| \ll 1$ , and  $D_{ii}(z) = B|z|^2$  for  $|z| \ll \ell_\nu$ . Here  $\alpha \in (0, 1]$  measures the space Hölder regularity of the field in the inertial range, and  $\ell_\nu$  is the dissipative scale. The infinite Reynolds number limit corresponds to  $\ell_\nu \rightarrow 0$  as  $\nu \rightarrow 0$ . See e.g. [Kup03], [Gaw08], [DE17], (2.26)–(2.27)].

(even in the presence of boundaries). We refer to [FGV01, Kup03, Gaw08, DE17] for excellent discussions about the Kraichnan model.

The main drawback of this model stems from the white-noise temporal correlation of the vector field  $\mathbf{b}(t, x)$ , which is indeed so rough that it is probably responsible for the anomalous diffusivity. At the experimental level, a consequence of this roughness was already noted by Sreenivasan and Schumacher [SS10]: there are several differences between the predictions of the Kraichnan model and the behavior of a passive scalar in Navier-Stokes turbulence. At the mathematical level, the white-noise temporal correlation allows for a certain explicit and exact computation of the statistics of the solution. Namely, one may obtain closed expressions for the correlation functions of the scalar  $\theta^\kappa$ ; therein, the assumed white in time correlation structure of the velocity field plays a crucial role. As a consequence, the “exact analysis” developed for the Kraichnan model is not robust, and it did not allow the fluid dynamics community to build sturdy tools for understanding the energy cascade in Navier-Stokes turbulence.

Nonetheless, as noted by Majda and Kramer in [MK99], exactly solvable models provide excellent test problems for assessing the strengths and weaknesses of approximate closure theories in turbulence. Besides the Kraichnan model discussed here, an exact mathematical analysis of diffusion (enhancement and anomalies) is also available for the “Simple Shear Models” of Avellaneda and Majda [AM91, AM90, AM92], which generalize an earlier model of Kubo [Kub63]. These examples emphasize how randomly fluctuating velocity fields act as effective diffusion processes, on large scales and long times. The vector fields in [AM91, AM90, AM92] are of a shear flow type  $\mathbf{b}(t, x) = (w(t), v(t, x_1))$ , where the spatially uniform sweeping component  $w(t)$  is taken as a stationary random process with possibly nonzero mean, and the shearing component  $v(x_1, t)$  is taken as a homogeneous and stationary, mean zero random field, whose statistics can be fine tuned to match the statistically stationary turbulent flows. Using exactly solvable renormalization group theories and Lagrangian renormalized perturbation theories (available for these simple shear flows), Avellaneda and Majda are able to identify several distinct regimes, as indexed by the mean of  $w$ , the strength of the infrared divergence in  $v$ , and the decorrelation time of long-wave portions of the statistical velocity spectrum. Note however that anomalous diffusion (1.4), is *not* available in the “Simple Shear Models” of [AM91, AM90, AM92].

**Singularly-focusing alternating shear flows.** To the best of our knowledge, the first example of a deterministic vector field  $\mathbf{b}(t, x)$ , for which the anomalous dissipation of scalar variance (1.4) is established rigorously, was recently constructed by Drivas, Elgindi, Iyer, and Jeong [DEIJ22]. In [DEIJ22, Theorem 1], it is shown that for any  $\alpha \in [0, 1)$  and  $d \geq 2$ , there exists a vector field  $\mathbf{b} \in L^1([0, 1]; C^\alpha(\mathbb{T}^d)) \cap L^\infty([0, 1]; L^\infty(\mathbb{T}^d))$ , such that the following holds:  $\mathbf{b}(t, \cdot)$  is smooth for any  $t < 1$ ; for any initial data with  $\theta_0 \in H^2$  which is sufficiently close (in  $L^2$ ) to a an eigenfunction of the Laplacian, anomalous diffusion (1.4) holds for some  $\varrho \in (0, 1)$ ; and the scalar field  $\theta^\kappa$  remains uniformly bounded in  $L^\infty([0, 1]; L^\infty(\mathbb{T}^d))$  as  $\kappa \rightarrow 0$ . The above result is sharp in the sense that if<sup>3</sup>  $\mathbf{b} \in L_t^1 W_x^{1, \infty}$  (corresponding to  $\alpha = 1$ ), then trivially one has  $\lim_{\kappa \rightarrow 0^+} \kappa \langle |\nabla \theta^\kappa|^2 \rangle = 0$ , as discussed in the first paragraph of Section 1.1.

In essence, the construction of the vector field  $\mathbf{b}(t, x)$  in [DEIJ22] alternates shear flows with

---

<sup>3</sup>The Lipschitz regularity may be replaced with merely the integrability of  $\nabla \mathbf{b}$ . Indeed, it follows from the Di Perna-Lions theory [DL89] that as soon as  $\mathbf{b} \in L_t^1 W_x^{1, 1}$  is divergence free, all bounded weak solutions of the transport equation  $\partial_t \theta + \operatorname{div}(\mathbf{b} \theta) = 0$  are renormalized, and thus they conserve the energy  $\frac{1}{2} \|\theta(t, \cdot)\|_{L^2}^2$ . See also the work of Ambrosio [Amb04] for  $\mathbf{b} \in L_t^1 \operatorname{BV}_x$ , divergence free. Then, as  $\kappa \rightarrow 0^+$  the a priori (subsequential) weak convergence of  $\theta^\kappa$  to a weak solution  $\theta$  of the transport equation, is in fact strong (due to the energy balance (1.5) and lower-semicontinuity), implying that there is no dissipation anomaly.

stream function<sup>4</sup>  $\sin(2^{(1+\alpha)j} x_{1+j \bmod d})$ , on intervals  $[t_{j-1}, t_j]$ , for  $\{t_j = 1 - 2^{-j} : j \geq 1\}$ . This construction is on the one hand inspired by the earlier work of Pierrehumbert [Pie94], who proposed an alternating shear flow of a single frequency, but with random i.i.d. phase shifts, to construct a “universal mixer” for the transport equation.<sup>5</sup> On the other hand, the idea of a quasi self-similar evolution on  $[t_{j-1}, t_j] \times \mathbb{T}^d$  which singularly focuses as  $j \rightarrow \infty$  all the “action” towards the final time slice  $\{t = 1\} \times \mathbb{T}^d$ —where all the anomalous dissipation of scalar variance occurs—is inspired by earlier works of Aizenman [Aiz78] and Depauw [Dep03] concerning the uniqueness of the transport equation below, and Alberti, Bianchini, and Crippa [ABC14] respectively Alberti, Crippa, and Mazzucato [ACM19a, ACM19b] regarding mixing for the transport equation.<sup>6</sup> With  $\mathbf{b}(t, x)$  constructed as such, the proof of [DEIJ22] hinges on comparing the family of solutions  $\{\theta^\kappa\}_{\kappa>0}$  to a solution  $\theta^0$  of the transport equation ( $\kappa = 0$ ) which satisfies  $\lim_{t \rightarrow 1^-} \|\nabla \theta^0\|_{L^2((0,t) \times \mathbb{T}^d)} = +\infty$  and for which a significant amount of energy travels to higher and higher frequencies as  $t \rightarrow 1^-$ , either as inviscid mixing or as a balanced growth of Sobolev norms, resulting in a lack of compactness at time  $t = 1$ ; see the abstract criterion for anomalous dissipation in [DEIJ22, Corollary 1.5].

Alternating shear flows which focus in a singular and quasi self-similar way onto a final time slice have also been recently considered by Brue and De Lellis [BDL22], Colombo, Crippa, and Sorella [CCS22], and jointly in [BCC<sup>+</sup>22], to give examples of anomalous dissipation of energy for solutions of the *forced* 3D Navier-Stokes equations [BDL22, BCC<sup>+</sup>22], and to establish anomalous diffusion for the drift-diffusion equation together with *uniform-in-diffusivity Hölder regularity* for the associated passive scalar. At the core of all these works is the anomalous dissipation of scalar variance for the drift-diffusion equation (1.4).

Indeed, it is well-known that for  $2\frac{1}{2}$ -dimensional solutions of the 3D Navier-Stokes equations the vertical component of the flow satisfies the linear advection-diffusion equation (1.1). More precisely, if  $u_H = (u_1, u_2)(x_1, x_2, t) : \mathbb{T}^2 \times \mathbb{R} \rightarrow \mathbb{R}^2$  and  $u_3 = (x_1, x_2, t) : \mathbb{T}^2 \times \mathbb{R} \rightarrow \mathbb{R}$  satisfy  $\partial_t u_H + (u_H \cdot \nabla_H) u_H + \nabla_H p - \nu \Delta_H u_H = f_H$ , respectively  $\partial_t u_3 + (u_H \cdot \nabla_H) u_3 - \nu \Delta_H u_3 = 0$ , where  $f_H$  is a horizontal body force,  $p$  is a scalar pressure ensuring  $\nabla_H \cdot u_H = 0$ , and we denote “horizontal” differential operators by  $\nabla_H = (\partial_{x_1}, \partial_{x_2})$  and  $\Delta_H = \partial_{x_1 x_1} + \partial_{x_2 x_2}$ , then the vector field  $u = (u_H, u_3)$  solves the 3D Navier-Stokes equations with viscosity  $\nu$ , pressure  $p$ , and body force  $(f_H, 0)$ . Then, inspired by the constructions in [Aiz78, Dep03, ABC14, ACM19a, ACM19b] the papers [BDL22, CCS22, BCC<sup>+</sup>22] construct both initial data for the scalar  $u_3$  (essentially a  $\pm 1$  checkerboard at unit scale) and a two-dimensional vector field  $u_H$ —which is essentially a sequence of alternating shear flows which are quasi self-similar on intervals of the type  $[t_{j-1}, t_j]$  with amplitudes  $a_j$  and frequencies  $\lambda_j$ , where  $t_j \rightarrow 0^+$ ,  $a_j, \lambda_j \rightarrow \infty$  as  $j \rightarrow \infty$ —such that the the inviscid transport equation  $\partial_t u_3 + (u_H \cdot \nabla_H) u_3 = 0$  mixes perfectly as  $t \rightarrow 1^-$ , i.e.  $u_3(t, \cdot) \rightarrow 0$  as  $t \rightarrow 1^-$ . To incorporate the effect of a vanishing sequence of diffusions  $\nu_j \rightarrow 0^+$  as  $j \rightarrow \infty$ , these

<sup>4</sup>The sinusoidal shear velocity profiles are replaced by a smoothed sawtooth function, which makes the computations easier, and in fact almost explicit.

<sup>5</sup>The proof that the Pierrehumbert construction indeed an universal exponential mixer was recently obtained by Blumenthal, Coti Zelati, and Gvalani [BCZG22], using a random dynamical systems based perspective.

<sup>6</sup>The deterministic theory of *mixing* for the linear transport ( $\kappa = 0$ ) and of *enhancement of diffusion* for the drift-diffusion ( $\kappa > 0$ ) equation (1.1) is too vast to review here. Usually these theories consider vector fields  $\mathbf{b}$  whose regularity is at least  $L_t^1 W_x^{1,p}$  for  $p \geq 1$ , so that the Di Perna-Lions theory applies to bounded solutions of the scalar linear transport. The questions typically asked are: when  $\kappa = 0$ , to describe the decreasing function  $\varrho(t)$  and the timescale  $t_0$  such that  $\|\theta^0(t, \cdot)\|_{H^{-1}} \leq \varrho(t - t_0) \|\theta^0(t_0, \cdot)\|_{H^1}$ , see e.g. [CDL08, IKX14, Sei13, EZ19]. In other works, the loss of regularity and nonuniqueness of weak solutions to the continuity equation is discussed [ABC14, Jab16, YZ17, ACM19a, ACM19b, CEIM22] and [MS18, BCDL21, CL21]. For  $\kappa > 0$ , it is well-known that some enhancement of diffusion takes place due to mixing properties of the underlying flow of  $\mathbf{b}$  [CKRZ08, BCZ17, FI19, CZDE20, CZD21]. For such diffusion enhancing flows, the challenge is to quantify the optimal rate  $r(\kappa) \gg \kappa$  and the timescale  $t_\kappa \ll \kappa^{-1}$  such that  $\|\theta^\kappa(t, \cdot)\|_{L^2}^2 \leq C e^{-r(\kappa)t} \|\theta^0\|_{L^2}^2$ , for all  $t \geq t_\kappa$  [CZDE20, CZD21, BN21, ELM23]. The dissipation anomaly considered in this paper is an extreme form of enhancement of diffusion, with  $r(\kappa) = \mathcal{O}(1)$  uniformly in  $\kappa$  as  $\kappa \rightarrow 0$ .

authors smooth out the aforementioned vector field at a specific  $(a_j, \lambda_j, t_j, \nu_j)$ -dependent scale, and then either appeal to the abstract criterion from [DEIJ22] or directly measure the variance of the associated stochastic process, to show that the drift-diffusion equation may be viewed as a perturbation of the transport equation, and hence exhibits anomalous diffusion. The term  $f_H$  is then just the remainder obtained by inserting the constructed vector  $u_H$  into the horizontal part of the 3D Navier-Stokes equations. A clever fine-tuning of the parameters  $(a_j, \lambda_j, t_j, \nu_j)$  in the construction attains both the uniform Hölder regularity of the sequence  $\{u^{\nu_j}\}_{j \geq 1}$  (in the full range strictly below  $L_t^1 W_x^{1,1}$ ), and fact that  $\limsup_{j \rightarrow \infty} \nu_j \|\nabla u^{\nu_j}\|_{L^2((0,1) \times \mathbb{T}^3)}^2 > 0$ . As in [DEIJ22], in these constructions the anomalous dissipation occurs only on the time slice  $\{t = 1\} \times \mathbb{T}^d$ . We also note that by adding an extra space dimension to replace time, Johansson and Sorella [JS23] have obtained similar results for the advection-diffusion equation in dimensions larger than 3, for a vector field which is *autonomous*; here, the quasi self-similar singular focusing is achieved on a “last space slice” instead of a “last time slice” (see also [Aiz78, Figure 3] for a closely related idea).

The main drawbacks of the aforementioned constructions of [DEIJ22] and of [BDL22, CCS22, BCC<sup>+</sup>22] are as follows: (i) all the energy that can be dissipated anomalously is dissipated at only one instant in time, (ii) the vector field  $\mathbf{b}(t, x)$  has only one active scale at each time  $t \in [0, 1]$ , (iii) the drift-diffusion equation is treated as a perturbation of the transport equation, and (iv) the vector field  $\mathbf{b}(t, x)$  and the initial datum  $\theta_0$  are not constructed independently of each other, and the diffusive anomaly is not proved for all smooth initial data.

Regarding point (i), we note that the existence of a single time (e.g.  $t = 1$  for [DEIJ22, BDL22, CCS22, BCC<sup>+</sup>22]) at which all of the anomalous diffusion occurs, is incompatible with the (statistical) *stationarity* of the turbulent vector fields, for which anomalous diffusion has been robustly observed in practice. In contrast, the vector field  $\mathbf{b}(t, x)$  which we construct in Theorem 1.1 does not distinguish any special times, and for  $t_1, t_2 \in [0, 1]$  chosen at random,  $\mathbf{b}(t_1, \cdot)$  and  $\mathbf{b}(t_2, \cdot)$  have the same regularity, are macroscopically undistinguishable. This means in particular that our vector field does not quasi self-similarly focus the dynamics onto a single time slice, leading us to point (ii). In the previous examples of anomalous diffusion [DEIJ22, BDL22, CCS22, BCC<sup>+</sup>22] at each instance of time  $t \in [0, 1]$  only one shear flow is active (at a suitable spatial frequency), which in turn necessitates singular focusing in time for the passive scalar to witness infinitesimally small scales in  $\mathbf{b}(t, x)$ . This picture is inconsistent with the observed power spectra of turbulent flows in statistical equilibrium [Fri95]. The vector field  $\mathbf{b}(t, x)$  from Theorem 1.1 does not have this property: at a.e.  $t \in [0, 1]$  the vector field  $\mathbf{b}(t, \cdot)$  contains infinitely many shear flows of diverging frequencies, which are twisted by the Lagrangian flows induced by the sum of the flows at all scales “above” that of the shear being considered. At first sight, one may think that this “feature” of  $\mathbf{b}(t, x)$  comes with a “bug”: the underlying transport equation is severely ill-posed, leading us to point (iii). At the heart of the proofs in [DEIJ22, BDL22, CCS22, BCC<sup>+</sup>22], the transport equation does the heavy lifting, in a quasi self-similar fashion as  $t \rightarrow 1^-$ . In a sense, it is shown that the non-diffusive picture is stable in  $L^2$  under diffusive perturbations. Our work presents a fundamental difference, as we do not view (1.1) as a perturbation of the transport equation  $(\partial_t + \mathbf{b} \cdot \nabla)\theta = 0$ . In fact, the diffusion is used in a fundamental way in the proof (see Section 1.2). At each scale larger than the smallest active scale (determined by  $\kappa$ ) the advection part of the operator is in balance with a renormalized diffusion operator. A welcome consequence of this perspective and of this proof strategy is that in our analysis the vector field  $\mathbf{b}(t, x)$  and the initial data  $\theta_0$  are independent of each other, with (1.4) holding for every  $\theta_0 \in \dot{H}^1(\mathbb{T}^d)$ . This “universality” was however not present in any of the earlier works on this subject, as mentioned in point (iv) above. In [BDL22, CCS22, BCC<sup>+</sup>22] the main results establish the existence of both a vector field  $\mathbf{b}(t, x)$  and of an initial datum  $\theta_0$  (a  $\pm 1$  checkerboard) for which (1.4) holds, while in [DEIJ22, Theorem 1] the initial datum needs to

be sufficiently close (with respect to the  $H^2$  topology) to an eigenfunction of the Laplacian on  $\mathbb{T}^d$ .<sup>7</sup> This is of course not the physically motivated problem since the turbulent vector field  $\mathbf{b}(t, x)$  should be given in advance (as a solution of, say, 3D Navier-Stokes), and then the passive scalar is to be advected and diffused in this flow. The reason why Theorem 1.1 yields anomalous diffusion for all  $H^1$  initial data of zero mean is not the construction of the vector field  $\mathbf{b}(t, x)$  per se, it is the proof strategy, which shows that the quantity  $\kappa \|\nabla \theta^\kappa\|_{L^2((0,1) \times \mathbb{T}^d)}^2$  is close (in a  $\kappa$ -independent sense) to the rate of diffusion experienced by (essentially) a heat equation with the same initial datum, and unit-size diffusivity coefficient.

## 1.2 An outline of the proof: fractal homogenization

We present the proof of Theorem 1.1 only in dimension  $d = 2$  rather than a general dimension  $d \geq 2$  for convenience and readability. The argument in higher dimensions has only notational differences.

As mentioned above, the proof of Theorem 1.1 is based on the idea that anomalous diffusivity is the consequence of a “homogenization cascade” of “eddy diffusivities,” which goes from small scales to large scales. We think of each homogenization step as modifying the equation by removing the fastest wiggles in the vector field and—due to the enhancement of diffusivity caused by these wiggles—increasing the diffusivity parameter  $\kappa$ . The “effective diffusivities” thereby increase as we zoom out to larger scales, until finally, at the macroscopic scale, the vector field has no remaining wiggles and the effective diffusivity is of order one. This strategy, which is a renormalization group-type approach, has a very long history dating back to the 19th century (see [Fri95, Chapter 9]).

In this subsection, we will give a complete overview of the main ideas behind the construction of the vector field  $\mathbf{b}(t, x)$  and the proof of anomalous dissipation of scalar variance. The full proof is very lengthy, as the justifications of many of the intuitions here require long computations and many estimates.

**Advection-enhanced diffusion and homogenization.** We briefly review the phenomenon of advection-enhanced diffusion, from the point of view of classical homogenization. Consider a  $\mathbb{Z} \times \mathbb{Z}^d$ -periodic, mean-zero, incompressible vector field  $\mathbf{u}(t, x)$  and the advection-diffusion equation

$$\partial_t \theta_\varepsilon - \kappa \Delta \theta_\varepsilon + \frac{1}{\varepsilon} \mathbf{u}(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon}) \cdot \nabla \theta_\varepsilon = 0. \quad (1.9)$$

The advection term may be expressed as a second-order term:

$$\frac{1}{\varepsilon} \mathbf{u}(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon}) \cdot \nabla \theta_\varepsilon = -\nabla \cdot (\mathbf{s}(\frac{\cdot}{\varepsilon^2}, \frac{\cdot}{\varepsilon}) \nabla \theta_\varepsilon),$$

where  $\mathbf{s}$  is a *stream matrix* for  $\mathbf{u}$ , that is, an anti-symmetric matrix such that  $-\nabla \cdot \mathbf{s} = \mathbf{u}$ . This allows us to write (1.9) as

$$\partial_t \theta_\varepsilon - \nabla \cdot (\kappa \mathbf{I}_d + \mathbf{s}(\frac{\cdot}{\varepsilon^2}, \frac{\cdot}{\varepsilon})) \nabla \theta_\varepsilon = 0. \quad (1.10)$$

Here there are two scales: the small scale  $\varepsilon > 0$  on which the stream matrix oscillates, and the macroscopic scale which is of order one. Classical homogenization theory says that (1.10) homogenizes to the *effective equation*

$$\partial_t \bar{\theta} - \nabla \cdot \bar{\mathbf{a}} \nabla \bar{\theta} = 0,$$

---

<sup>7</sup>This limitation also applies to the constructions based on intermittent convex integration schemes [BV20] applied to the transport and drift diffusion equations e.g. in [MS18, BCDL21, CL21, PS21]: all of these construct the vector field at the same time as the scalar.

in the sense that, roughly speaking, solutions of the former converge in  $L^2$ , as  $\varepsilon \rightarrow 0$ , to those of the latter. The effective diffusion matrix  $\bar{\mathbf{a}}$  is given by the formula

$$\bar{\mathbf{a}}e = \langle\langle (\kappa\mathbf{I}_d + \mathbf{s})(e + \nabla\chi_e) \rangle\rangle, \quad e \in \mathbb{R}^d,$$

where  $\langle\langle \cdot \rangle\rangle$  denotes the (space-time) average of a  $\mathbb{Z} \times \mathbb{Z}^d$ -periodic function and  $\chi_e$  is the *corrector* with slope  $e$ , that is, the unique periodic (in space and time), mean-zero solution of the cell problem

$$\partial_t\chi_e - \nabla \cdot (\kappa\mathbf{I}_d + \mathbf{s})(e + \nabla\chi_e) = 0.$$

The symmetric part of  $\bar{\mathbf{a}}$  is given by

$$\frac{1}{2}(\bar{\mathbf{a}} + \bar{\mathbf{a}}^t)_{ij} = \kappa\delta_{ij} + \kappa\langle\langle \nabla\chi_{e_i} \cdot \nabla\chi_{e_j} \rangle\rangle.$$

The second term is positive (in the ordering of nonnegative definite matrices), and therefore the symmetric part of  $\bar{\mathbf{a}}$  is larger than the original diffusion matrix  $\kappa\mathbf{I}_d$ . This effect is called the enhancement of diffusivity due to advection.

The enhancement of diffusivity from the point of view of homogenization has been well-studied over the past four decades. There are too many works to cite here, so we refer the reader to [FP94, MK99] and the references therein. The proposal to use homogenization methods to turbulence models has however received a great deal of skepticism, primarily due to the lack of asymptotic scale separation. The homogenization limit requires sending the parameter  $\varepsilon$ , representing the ratio of the two scales, to zero. Such criticisms can be found in [Fri95, page 225] and [MK99, page 304].

Indeed, the vector field  $\mathbf{b}(t, x)$  we will construct will have certain active scales and the ratio of any pair of these active scales is fixed and not parametrized by a parameter being sent to zero. Moreover, we have infinitely many active scales, and not only two as in the simple setup described above. These issues pose serious analytic challenges, and we will address them using *quantitative* homogenization methods. Rather than reason in terms of asymptotic limits, we need to precisely quantify the length scales and time scales on which homogenization occurs. We will next consider this question in the context of a simple shear flow.

**Homogenization of a simple shear flow.** The vector field  $\mathbf{b}(t, x)$  will have a fractal-like structure, and so we need to introduce the “base” of the fractal, that is, the pattern which links two different scales and will be repeated infinitely many times. This role will be served by a simple alternating shear flow.

Given parameters  $a, \varepsilon > 0$ , consider the simple time-independent shear flow  $\mathbf{u}(x)$  defined by

$$\mathbf{u}(x) = \begin{pmatrix} 0 \\ 2\pi a \varepsilon \cos(\frac{2\pi x_1}{\varepsilon}) \end{pmatrix}.$$

The length scale on which the shear flow varies is  $\varepsilon$ , and the parameter  $a$  represents the size of the Lipschitz norm of the vector field. The stream function for  $\mathbf{u}$  is  $\psi_\varepsilon(x) = a\varepsilon^2 \sin(\frac{2\pi x_1}{\varepsilon})$ ; in other words,  $\mathbf{u} = \nabla^\perp \psi_\varepsilon$ . We stress here that  $\varepsilon$  is not a parameter to be sent to zero, it just represents the inverse wave number of  $\mathbf{u}$ .

The equation for a passive scalar  $\theta$  advected by  $\mathbf{u}$  with diffusivity  $\kappa > 0$  can be written as

$$\partial_t\theta - \nabla \cdot (\mathbf{K}^\varepsilon(x)\nabla\theta) = 0 \tag{1.11}$$

where  $\mathbf{K}^\varepsilon$  is the non-symmetric matrix

$$\mathbf{K}^\varepsilon(x) = (\kappa \mathbf{I}_2 + \psi_\varepsilon(x) \sigma) = \begin{pmatrix} \kappa & -a\varepsilon^2 \sin(\frac{2\pi x_1}{\varepsilon}) \\ a\varepsilon^2 \sin(\frac{2\pi x_1}{\varepsilon}) & \kappa \end{pmatrix}, \quad (1.12)$$

and  $\sigma$  is defined in (1.20), below. Due to homogenization, we expect that (1.11) should be close, on large enough scales, to its effective equation

$$\partial_t \theta - \nabla \cdot \bar{\mathbf{K}}_2 \nabla \theta = 0,$$

where the effective diffusivity matrix  $\bar{\mathbf{K}}_2$  in this case can be computed explicitly. It is:

$$\bar{\mathbf{K}}_2 = \begin{pmatrix} \kappa & 0 \\ 0 & \kappa + \frac{a^2 \varepsilon^4}{\kappa} \end{pmatrix}. \quad (1.13)$$

We again stress that we are not sending  $\varepsilon \rightarrow 0$  here, the homogenization is with respect to a large-scale limit.

The formula (1.13) was derived more than 70 years ago by Taylor [Tay53] and is sometimes called the *Taylor dispersion formula* (specialized to the shear flow); see also [Ari56]. As he showed, homogenization is observed on length scales much larger than  $a\varepsilon^3\kappa^{-1}$  and time scales much larger than  $\varepsilon^2\kappa^{-1}$ . This can be proved analytically from estimates on the correctors, which can be computed explicitly in this case.

There is another way to think about this, in terms of the particle trajectories. The diffusion  $Y_t$  process corresponding to (1.11) satisfies the SDE

$$dY_t = \mathbf{u}(t, Y_t) dt + \sqrt{2\kappa} dW_t. \quad (1.14)$$

The particle evolving according to these dynamics will move with speed of order  $a\varepsilon$  in the  $x_2$  direction, changing its direction (up or down) and its magnitude on time scales of order  $\varepsilon^2\kappa^{-1}$ , which is the time it takes the diffusion to alter its  $x_1$  coordinate on the order of  $\varepsilon$ . The vector field has typical size  $a\varepsilon$ , therefore in this time the particle will have travelled a distance of order  $a\varepsilon \cdot \varepsilon^2\kappa^{-1} = a\varepsilon^3\kappa^{-1}$ . If we zoom out and observe the motion of the particle on length scales much larger than  $a\varepsilon^3\kappa^{-1}$  and time scales much larger than  $\varepsilon^2\kappa^{-1}$ , then what we see (roughly) is that the  $x_2$  coordinate of the particle is performing a random walk with steps of size  $a\varepsilon^3\kappa^{-1}$ , with  $\varepsilon^2\kappa^{-1}$  units of time between steps. This leads to a diffusivity in the  $x_2$  direction of order

$$\frac{(a\varepsilon^3\kappa^{-1})^2}{\varepsilon^2\kappa^{-1}} = \frac{a^2\varepsilon^4}{\kappa}.$$

Of course, this diffusive effect caused by the advection should be in addition to the molecular diffusion, so we expect to find an effective diffusion of order

$$\text{effective diffusivity in } x_2 \text{ direction} = O\left(\kappa + \frac{a^2\varepsilon^4}{\kappa}\right).$$

This rough intuition is in agreement the more precise formula (1.13).

The dimensionless quantity  $a\varepsilon^2\kappa^{-1}$ , recognized as representing the (square root of the) ellipticity contrast in the matrix  $\mathbf{K}^\varepsilon$  defined in (1.11), is a measure of the strength of the shear flow term relative to the molecular diffusion. It determines the multiple of the small scale  $\varepsilon$  on which homogenization occurs.

Now consider a vector field which alternates between shear flows in the  $x_1$  direction and shear flows in the  $x_2$  direction with frequency  $\tau^{-1}$ :

$$\mathbf{v}_{\varepsilon,a,\tau}(t,x) := 2\pi a\varepsilon \sum_{k \in \mathbb{Z}} \mathbf{1}_{[k\tau,(k+1)\tau)}(t) \left( \begin{pmatrix} 0 \\ \cos(\frac{2\pi x_1}{\varepsilon}) \end{pmatrix} \mathbf{1}_{\{k \in 2\mathbb{Z}\}} + \begin{pmatrix} -\cos(\frac{2\pi x_2}{\varepsilon}) \\ 0 \end{pmatrix} \mathbf{1}_{\{k \in 2\mathbb{Z}+1\}} \right). \quad (1.15)$$

If we require that  $\tau \gg \varepsilon^2 \kappa^{-1}$ , so that the shear flows have enough time to homogenize, then the corresponding advection-diffusion equation homogenizes to the average of  $\bar{\mathbf{K}}_2$  and the analogous matrix  $\bar{\mathbf{K}}_1$  with the diagonal entries swapped, which is conveniently isotropic. We find that, on length scales much larger than  $a\varepsilon^3 \kappa^{-1}$  and time scales much larger than  $\tau$ , the equation with alternating shear flows will homogenize to

$$\partial_t \theta - \bar{\kappa} \Delta \theta = 0,$$

where the effective diffusivity is given by Taylor's formula,  $\bar{\kappa} = (1 + \frac{a^2 \varepsilon^4}{2\kappa^2})\kappa$ .

**The construction of the multiscale vector field  $\mathbf{b}(t,x)$ .** The above discussion suggests an idea for setting up a “homogenization cascade” by constructing a vector field  $\mathbf{b}(t,x)$  with many copies of the alternating shear flows on different scales. We look for a decreasing sequence of length scales  $\varepsilon_m \rightarrow 0$ , of time scales  $\tau_m \rightarrow 0$ , and of diffusivities  $\kappa_m \rightarrow 0$  and an increasing sequence of parameters  $a_m \rightarrow \infty$  which satisfy the following relations:

$$\begin{cases} \kappa_{m-1} = \left(1 + \frac{a_m^2 \varepsilon_m^4}{2\kappa_m^2}\right) \kappa_m, \\ a_m = \varepsilon_m^{\alpha-1}, \\ \tau_m \gg \frac{\varepsilon_m^2}{\kappa_m}, \\ \varepsilon_{m-1} \gg \left(\frac{a_m \varepsilon_m^2}{\kappa_m}\right) \varepsilon_m. \end{cases} \quad (1.16)$$

The last condition is to ensure that the wiggles we put in the vector field at scale  $\varepsilon_{m-1}$  do not interfere with the homogenization of those at scale  $\varepsilon_m$ . The condition on  $a_m$  is because we want the vector field to be Hölder continuous with exponent  $\alpha \in (0, 1)$ . We would then like to define a vector field  $\mathbf{b}(t,x)$  in a recursive way, by adding shear flows at each scale  $\varepsilon_m$ , roughly as follows: set  $\mathbf{b}_0 = 0$ , and then define

$$\mathbf{b}_m(t,x) := \mathbf{b}_{m-1}(t,x) + \mathbf{v}_{\varepsilon_m, a_m, \tau_m}(t,x).$$

The idea is that the vector field  $\mathbf{b}_{m-1}$  is “macroscopic” from the point of view of  $\mathbf{v}_{\varepsilon_m, a_m, \tau_m}$ , which will homogenize before spatial or temporal variations in  $\mathbf{b}_{m-1}$  are noticed. We will then define  $\mathbf{b} := \lim_{m \rightarrow \infty} \mathbf{b}_m$ . Note that this limit makes sense, due to the fact that the supremum of  $|\mathbf{v}_{\varepsilon_m, a_m, \tau_m}|$  is of order  $a_m \varepsilon_m = \varepsilon_m^\alpha$ , which is small and can be summed up (since the scales will be at least geometrically separated). The hope is then that we have set up the parameters in such a way that the advection-diffusion equation with diffusivity  $\kappa_m$  and vector field  $\mathbf{b}_m$  will homogenize to the one with diffusivity  $\kappa_{m-1}$  and vector field  $\mathbf{b}_{m-1}$ .

This however will not work without another crucial modification. The presence of the “macroscopic” vector field  $\mathbf{b}_{m-1}$  actually interferes with the homogenization of the wiggles represented by  $\mathbf{v}_{\varepsilon_m, a_m, \tau_m}$ . Indeed, this “macroscopic” term is essentially a constant from the point of view of the much faster field  $\mathbf{v}_{\varepsilon_m, a_m, \tau_m}$ , and a large constant drift added to a shear flow essentially

destroys the shear flow structure, with its very long streamlines, and consequently removes most of the enhancement of the diffusivity. In fact, a large constant background drift will destroy the enhancement of diffusivity of any time-independent flow: see Appendix A for details.

This presents a serious obstacle to building examples of continuous vector fields which exhibit of anomalous diffusion, since continuity implies that larger wave numbers should have larger amplitudes. The solution to this problem is to force the small scale shear flow  $\mathbf{v}_{\varepsilon_m, a_m, \tau_m}$  to be swept by the vector field  $\mathbf{b}_{m-1}$ , so that they appear to be stationary in the moving reference frame of a particle advected by  $\mathbf{b}_{m-1}$ . We do this by modifying the definition of  $\mathbf{b}_m$  as follows:

$$\mathbf{b}_m(t, x) := \mathbf{b}_{m-1}(t, x) + \mathbf{v}_{\varepsilon_m, a_m, \tau_m}(t, X_{m-1}^{-1}(t, x)),$$

where  $X_{m-1}$  is the flow for the vector field  $\mathbf{b}_{m-1}$ , that is, the solution of  $\partial_t X = \mathbf{b}_{m-1}(t, X)$ . If changed into Lagrangian coordinates, then the  $\mathbf{b}_{m-1}$  term would disappear, and the Laplacian term would only be slightly distorted. In this way, the vector field  $\mathbf{v}_{\varepsilon_m, a_m, \tau_m}$  can be homogenized without disturbing  $\mathbf{b}_{m-1}$ . This “self-advection” property—arising here naturally from the renormalization perspective as a way to gain a sufficient enhancement of diffusivity between two scales—is a property that real fluids have.<sup>8</sup>

This introduces a new complexity to our construction, because the inverse flows must be renewed on a time scale which is much less than the inverse of the Lipschitz constant of  $\mathbf{b}_{m-1}$ , which is of order  $a_m$ . Otherwise the distortion due to the flows becomes intractable. Therefore we modify our definition again by introducing  $\tau_m'' > \tau$  and defining

$$\mathbf{b}_m(t, x) := \mathbf{b}_{m-1}(t, x) + \sum_{l \in \mathbb{Z}} \mathbf{1}_{[l\tau_m'', (l+1)\tau_m'']}(t) \mathbf{v}_{\varepsilon_m, a_m, \tau_m}(t, X_{m-1}^{-1}(t, x, l\tau_m'')), \quad (1.17)$$

where  $X_{m-1}(t, x, s)$  is the flow for  $\mathbf{b}_{m-1}$  with  $X_{m-1}(s, x, s) = x$ . This gives us two new constraints:

$$\tau_m \ll \tau_m'' \ll a_{m-1}^{-1}.$$

The second constraint ensures that we have good estimates on the difference between our flows and the identity matrix. The first constraint  $\tau_m'' \gg \tau_m$  is needed because the periodic renewal of the inverse flows has caused new periodic wiggles (in time) to appear in our vector field, and these must also be homogenized! We need to make sure that the time scale of these wiggles does not interfere with the homogenization problem for  $\mathbf{v}_{\varepsilon_m, a_m, \tau_m}$ .

We next try to see if we can choose the parameters to satisfy all of the constraints. First, in order to have  $\kappa_{m-1}$  much larger than  $\kappa_m$ , we need that

$$\frac{a_m \varepsilon_m^2}{\kappa_m} \gg 1.$$

This suggests that we should try to pick all the parameters so that this number is a negative power, say  $-\gamma < 0$ , of  $\varepsilon_m$ :

$$\frac{a_m \varepsilon_m^2}{\kappa_m} \simeq \varepsilon_m^{-\gamma}.$$

<sup>8</sup>This property is not shared by many other models of synthetic turbulence, to our knowledge. See [TD05, EB13] for a discussion of this point, and the apparent disparities between models of synthetic turbulence and real turbulent flows. Summarizing [TD05], the authors of [EB13] write “The key point is that large-scale eddies in real turbulence advect both particles and smaller scale eddies, while large-scale eddies in synthetic turbulence advect only particle pairs and not smaller eddies.” We remark that a related difficulty was faced by Kraichnan in his attempts to build his “DIA” model of turbulence: see [EF11, Section II.A], in particular the discussion of random Galilean invariance.

In other words, we have now chosen  $\kappa_m = a_m \varepsilon_m^{2+\gamma} = \varepsilon_m^{1+\alpha+\gamma}$ . The constraints for  $\tau_m$  reduce to  $\varepsilon_m^2 \kappa_m^{-1} \ll a_{m-1}^{-1}$ , which can be written in terms of the  $\varepsilon_m$ 's as

$$\varepsilon_m^{1-\alpha-\gamma} \ll \varepsilon_{m-1}^{1-\alpha}.$$

This is a sharper constraint than the one for the  $\varepsilon_m$ 's in (1.16), so it remains to check if it is compatible with the recurrence relation for  $\kappa_m$ . This will be the case if and only if

$$\frac{2(1-\alpha)}{1+\alpha+\gamma} > 1.$$

We can therefore pick an appropriate  $\gamma > 0$  if and only if  $2(1-\alpha)/(1+\alpha) > 1$ , which is equivalent to  $\alpha < 1/3$ . This is the reason for the restriction on  $\alpha$  in the statement of Theorem 1.1.

The scales  $\varepsilon_m$  are decreasing supergeometrically, and thus so are the diffusivities  $\kappa_m$ . The recurrence for  $\kappa_m$  in (1.16) is very sensitive to the initial choice of  $\kappa$ , and for this reason it must be chosen to be within a factor of two of  $\varepsilon_N^{1+\alpha+\gamma}$ , for some  $N$ . Otherwise the diffusivities will oscillate between very large and very small numbers, and we will lose control of our homogenization estimates.

Note that the exponents  $1 + \alpha + \gamma$  and  $(1+\alpha+\gamma)/(1-\alpha)$  in scaling of the renormalized diffusivities,  $\kappa_m \simeq \varepsilon_m^{1+\alpha+\gamma} \simeq \tau_m^{\frac{1+\alpha+\gamma}{1-\alpha}}$ , tend to  $4/3$  and  $2$ , respectively, as  $\alpha \rightarrow 1/3$ , which is in agreement with Richardson's  $4/3$  law. In fact, then the variance in the position  $Y_t$  of a particle trajectories at time  $t$  will indeed scale like  $t^{1+\frac{1+\alpha+\gamma}{1-\alpha}}$  (for small  $\kappa$ , well-chosen as explained above), and this exponent is close to  $3$  when  $\alpha$  is close to  $1/3$ , as predicted. Demonstrating this is outside the scope of the present paper, as it requires some uniform estimates for the passive scalar which will be the focus of a forthcoming paper [ARV]. These are analogous to large-scale regularity estimates in homogenization theory, adapted to the present situation of “fractal” homogenization. See below in Section 1.3 for more.

What is described above is a slight simplification of construction of the vector field  $\mathbf{b}(t, x)$  in Section 2. In the actual construction, the indicator functions of the time variable appearing in (1.15) and (1.17) are replaced by smooth approximations, so that the vector fields  $\mathbf{b}_m$  are smooth in time as well as space, and uniformly  $\alpha$  Hölder continuous in both variables. We also have “quiet” time intervals each time we switch the direction of the shear flows, which is convenient for technical reasons. We similarly arrange for the shear flows to pause on time intervals of length  $\tau'_m$  around any change of the inverse flows  $X_{m-1}^{-1}$ , where  $\tau_m \ll \tau'_m \ll \tau''_m$  is intermediate between the other two time scales.

**The homogenization step.** The reader is hopefully convinced that the vector field  $\mathbf{b}(t, x)$  whose construction we have outlined above is a good candidate for exhibiting anomalous diffusion.

However, analyzing the effect of the complicated fractal-like structure of the vector field  $\mathbf{b}(t, x)$  on the passive scalar  $\theta^\kappa$  is a challenge. Periodic homogenization is of course very well-understood, even if there are a large (but finite) number of well-separated scales (a topic referred to as *reiterated* homogenization). What is not well-understood is the case in which there are essentially infinitely many scales which are not well-separated. This is the situation we encounter here, because even if the ratio  $\varepsilon_{m-1}/\varepsilon_m$  between scales can be made arbitrarily large in our construction of  $\mathbf{b}(t, x)$ , once the vector field is constructed it is a fixed finite number.

This difficulty has not gone unnoticed. Indeed, the idea of renormalization group-type approach to anomalous diffusion for a passive scalar equation in which “eddy diffusivities” are successively

renormalized is described very clearly at a heuristic level in [Fri95, Section 9.6]. As explained there, this idea has been present since the 19th century, but it is the lack of clear scale separation that is the primary reason for the limited applicability of homogenization theory to passive scalar turbulence and the reason “why the concept of eddy viscosity has been regarded by some theoreticians of turbulence as (at best) a pedagogical device.” Similar remarks can be found in Majda and Kramer [MK99]. The present paper is the first work to our knowledge to address this difficulty in a fully rigorous way.<sup>9</sup>

Let  $\theta_m$  be the solution of the equation

$$\partial_t \theta_m - \kappa_m \Delta \theta_m + \mathbf{b}_m \cdot \nabla \theta_m = 0 \quad \text{in } (0, \infty) \times \mathbb{T}^2$$

with advecting vector field  $\mathbf{b}_m$ . As alluded to above, the main step in the proof of Theorem 1.1 is the demonstration that the equation for  $\theta_m$  homogenizes to the one for  $\theta_{m-1}$ . The precise statement is given below in Proposition 5.2, where one finds the estimate

$$\left| \frac{\kappa_m \|\nabla \theta_m\|_{L^2((0,1) \times \mathbb{T}^2)}^2}{\kappa_{m-1} \|\nabla \theta_{m-1}\|_{L^2((0,1) \times \mathbb{T}^2)}^2} - 1 \right| \leq C \varepsilon_{m-1}^\delta. \quad (1.18)$$

Here  $C$  is a constant depending only on  $\alpha$  and  $\delta > 0$  is an explicit exponent. An iteration of this estimate yields that the quantity  $\kappa_m \|\nabla \theta_m\|_{L^2((0,1) \times \mathbb{T}^2)}^2$  is nearly independent of  $m$ , up to an error which can be made very small, and in particular much smaller than  $\kappa_0 \|\nabla \theta_0\|_{L^2((0,1) \times \mathbb{T}^2)}^2$  which is of order one.

The basic idea of the proof of the homogenization step is simple and classical. We build an explicit ansatz for the solution of the equation for  $\theta_m$ , which we denote by  $\tilde{\theta}_m$ . This function is constructed explicitly using ingredients from the equation for  $\theta_{m-1}$ , so we have a good understanding of it—we know in particular that the difference  $\tilde{\theta}_m - \theta_{m-1}$  is small in  $L^2$ . We then plug  $\tilde{\theta}_m$  into the equation for  $\theta_m$  and carefully compute the error. If the error is sufficiently small, then we can deduce that the difference  $\theta_m - \tilde{\theta}_m$  is small from basic energy estimates.

The definition of the ansatz  $\tilde{\theta}_m$  can be found in (4.24). It is more complicated than usually expected for periodic homogenization. The classical two-scale ansatz consists of taking the periodic correctors  $\chi_e$  and attaching them to a solution of the macroscopic equation; in our setting, this suggests that we should define the two-scale ansatz by

$$\tilde{\theta}_m(t, x) := \theta_{m-1}(t, x) + \sum_{i=1}^d \varepsilon_m \chi_{e_i} \left( \frac{x}{\varepsilon_m} \right) \partial_{x_i} \theta_{m-1}(t, x),$$

where the  $\chi_{e_i}$  represent the correctors corresponding to the shear flows  $\mathbf{v}_{\varepsilon_m, a_m, \tau_m}$  on scale  $\varepsilon_m$  in the definition of  $\mathbf{b}_m$ . Keeping in mind that these shear flows are composed with the inverse flows  $X_{m-1}^{-1}$

---

<sup>9</sup>Here we are thinking of vector fields which are continuous. There have been previous works, such as [BAO02, KO02], which use homogenization methods to prove the superdiffusivity of stochastic processes advected by divergence-free vector fields. These papers consider a different scaling—there is no small diffusivity parameter  $\kappa$  being sent to zero—and the vector fields considered in these papers have many active scales, with the property that smaller scale wiggles (larger wave numbers) have much larger amplitudes. Note that the latter property is not consistent with continuity, if the vector field were to be rescaled (blown down). Indeed, viewed from the macroscopic scale, these vector fields will belong to negative regularity spaces. Building examples of such vector fields which exhibit superdiffusivity or anomalous diffusion is a much easier task, since one does not have the problem, mentioned above, of the low wave numbers killing the enhancement of the large wave numbers.

in the definition of  $\mathbf{b}_m$ , which are very slow compared to  $\mathbf{v}_{\varepsilon_m, a_m, \tau_m}$ , it is reasonable to compose these correctors with the inverse flows. Thus we should modify our ansatz to

$$\tilde{\theta}_m(t, x) := \theta_{m-1}(t, x) + \sum_{l \in \mathbb{Z}} \mathbf{1}_{[l\tau_m'', (l+1)\tau_m'']}(t) \sum_{i=1}^d \varepsilon_m(\chi_{e_i} \circ X_{m-1}^{-1}(t, \cdot, l\tau_m''))\left(\frac{x}{\varepsilon_m}\right) \partial_{x_i} \theta_{m-1}(t, x), \quad (1.19)$$

where as usual the indicator functions of time are actually replaced by a smoother approximation.

The actual ansatz we make is actually more complicated than (1.19), see (4.24) in Section 4.2 for the precise expression. Ultimately, our choice of  $\tilde{\theta}_m$  is justified by the estimate (5.29), which says  $\tilde{\theta}_m$  is sufficiently close to being a solution of the equation for  $\theta_m$  that we can deduce that the difference  $\theta_m - \tilde{\theta}_m$  is small. Obtaining the bound (5.29) turns out to be quite technical and much of the effort in Sections 3–5 is devoted to its proof.

### 1.3 Uniform estimates for the scalar

The predictions made by the phenomenological theories of scalar turbulence go of course much further than the anomalous dissipation of scalar variance; we refer the reader to [Fri95, SS00, War00, FGV01, DSY05, SS10] and references therein for a detailed account. For example, by drawing direct analogies with the Kolmogorov theory of fluid turbulence, Obukhov [Obu49] and Corrsin [Cor51] used scaling arguments to predict that if the vector field  $\mathbf{b}(t, x)$  represents a homogenous isotropic velocity field exhibiting K41 “monofractal” scaling in the inertial range, with exponent  $1/3$ , then the scalar field  $\theta^\kappa$  inherits this property—namely “monofractal” scaling of structure functions with exponent  $1/3$ —in the corresponding  $\kappa$ -dependent scalar inertial range. This scaling argument can be directly generalized to say that if the structure functions of  $\mathbf{b}(t, x)$  have monofractal scaling with exponent  $\alpha$ , then the structure functions of the scalar  $\theta^\kappa$  have monofractal scaling with exponent  $(1-\alpha)/2$ , for any  $0 < \alpha < 1$ , not just  $1/3$  (though this is the relevant exponent in fluid turbulence, at least in three spatial dimensions).

Just as with the Onsager conjecture, one may propose a *mathematical idealization*, corresponding to simultaneously diverging Reynolds and Péclet numbers, and postulate a dichotomy:<sup>10</sup> (i) if  $\mathbf{b} \in C_t^0 C_x^\alpha$  and if the solutions  $\{\theta^\kappa\}_{\kappa>0}$  of (1.1) are uniformly in  $\kappa$  bounded in  $C_t^0 C_x^{\alpha'}$  with  $\alpha' > (1-\alpha)/2$ , then  $\lim_{\kappa \rightarrow 0} \kappa \|\nabla \theta^\kappa\|_{L_t^2 L_x^2}^2 = 0$ , and (ii) there exists  $\mathbf{b} \in C_t^0 C_x^\alpha$  (presumably with  $\alpha < 1/3$ ), such that for all smooth initial conditions  $\theta_0$ , the solutions  $\{\theta^\kappa\}_{\kappa>0}$  of (1.1) are uniformly in  $\kappa$  bounded in  $C_t^0 C_x^{\alpha'}$  for any  $\alpha' < (1-\alpha)/2$ , and moreover  $\lim_{\kappa \rightarrow 0} \kappa \|\nabla \theta^\kappa\|_{L_t^2 L_x^2}^2 > 0$ .

Part (i) of this dichotomy is well-known and follows directly from the commutator estimate of Constantin, E, and Titi [CET94]. As stated above, part (ii) of this dichotomy is open. Theorem 1.1 does not address the Hölder regularity of the family  $\{\theta^\kappa\}_{\kappa>0}$ , only the anomalous dissipation of scalar variance. We note however that the paper [CCS22] establishes a version of this Hölder regularity, but the uniform in  $\kappa$  bounds for  $\|\theta^\kappa(t, \cdot)\|_{C_x^{\alpha'}}$ , in the full range  $\alpha' < (1-\alpha)/2$ , are only established for a particular initial datum  $\theta_0$ , and only in  $L^2$  with respect to the time variable.

In forthcoming joint work with Rowan [ARV], we will sharpen the statement of Theorems 1.1 in a number of ways. We show in [ARV] that with  $\mathbf{b}(t, x)$  as in Theorem 1.1, the advection-diffusion equation (1.1) regularizes the solutions up to  $C_t^0 C_x^{(1-\alpha)/2}$  along the subsequences  $\{\kappa_j\}$  exhibiting the diffusive anomaly (1.4). We show that for all  $t_0 \in (0, 1)$ , there exists a constant  $C_{t_0} = C_{t_0}(d, \alpha) > 0$  such that, for every mean-zero initial datum  $\theta_0 \in L^2(\mathbb{T}^d)$ , we have

$$\sup_{\kappa \in I} \|\theta^\kappa\|_{C^0([t_0, 1], C^{(1-\alpha)/2}(\mathbb{T}^d))} \leq C_{t_0} \|\theta_0\|_{L^2(\mathbb{T}^d)} < \infty,$$

<sup>10</sup>See also the discussions in [DEIJ22, Section 5] and [CCS22, Section 1].

where  $I$  is the interval of diffusivities defined in the paragraph below (1.6) above. This is achieved by complementing the argument in this paper with “large-scale regularity” techniques developed in quantitative homogenization theory. Another consequence of these estimates, which is obtained in [ARV], is that the diffusive anomaly  $\varrho$  in (1.4) is uniform in the initial datum: the dependence of  $\varrho(\theta_0)$  on  $\theta_0 \in L^2(\mathbb{T}^d)$  in Theorem 1.1 can be removed completely; in fact, we can take  $\varrho = 1/2$ . Moreover, by obtaining uniform-in- $\kappa$  estimates for the parabolic Green function associated to the drift-diffusion equation (1.1), we obtain in [ARV] estimates for the rate of separation of the squared distance between two realizations of the SDE process  $Y_t$  defined in (1.14), which are consistent with Richardson’s 4/3-law.

## 1.4 Notation

We denote the positive integers by  $\mathbb{N} = \{1, 2, \dots\}$ , and the non-negative integers by  $\mathbb{N}_0 = \mathbb{N} \cup \{0\} = \{0, 1, 2, \dots\}$ . We denote

$$\nabla^\perp f := \sigma \nabla f, \quad \text{where} \quad \sigma := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \quad (1.20)$$

We use the brackets  $\langle \cdot \rangle$  to denote the mean of a periodic function of space only (not time), that is,  $\langle \cdot \rangle = \int_{\mathbb{T}^d} (\cdot) dx$ . Averages of periodic functions in both space and time (or just time) are denoted by  $\langle\langle \cdot \rangle\rangle$ . We use  $\vee$  and  $\wedge$  to denote maximum and minimum operations, that is,  $a \vee b := \max\{a, b\}$  and  $a \wedge b := \min\{a, b\}$ . We denote the indicator function of a set  $A$  by  $\mathbf{1}_A$ . It is convenient to introduce, for every nonnegative integer  $n \in \mathbb{N}_0$  and  $f \in C^\infty(\mathbb{R}^2)$ ,

$$\llbracket f \rrbracket_{n,R} = \frac{(n+1)^2}{n!R^n} \sup_{|\alpha|=n} \|\partial^\alpha f\|_{L^\infty(\mathbb{R}^2)} \quad \text{and} \quad \llbracket f \rrbracket_R = \sup_{n \in \mathbb{N}_0} \llbracket f \rrbracket_{n,R}. \quad (1.21)$$

Estimates involving this seminorm are explored in Appendix B. The prefactor of  $(n+1)^2/n!$  is chosen so that the multiplicative property of the seminorm stated in Lemma B.1 is valid. Note that  $\llbracket \cdot \rrbracket_{n,R}$  and  $\llbracket \cdot \rrbracket_R$  are monotone decreasing with respect to  $R$ . For a Banach space  $X$  with norm  $\|\cdot\|_X$ , for  $n \in \mathbb{N}_0$ , and for a sufficiently smooth function  $f$ , it is convenient to denote the  $X$ -norm of the  $n$ th order symmetric tensor  $(\partial^\alpha f)_{|\alpha|=n}$  as

$$\|\nabla^n f\|_X := \max_{|\alpha|=n} \|\partial^\alpha f\|_X. \quad (1.22)$$

We allow constants  $C$  to vary from line to line. If we need to refer to a specific  $C$  from a particular displayed equation, we put the equation number in the subscript of  $C$ ; e.g.,  $C_{(2.65)}$  refers to the constant  $C$  on the right side of (2.65).

## 1.5 Outline of the paper

In Section 2, we construct the vector field  $\mathbf{b}(t, x)$  and give bounds on the derivatives of its approximations  $\mathbf{b}_m$ . We also study the regularity of the flows and inverse flows associated to  $\mathbf{b}_m$ . In Section 3, we build the correctors, define the sequence of renormalized diffusivities and other objects which are needed in the homogenization step. The ansatz  $\tilde{\theta}_m$  is introduced in Section 4, where some important estimates are also proved, using ingredients from Section 3. The main part of the argument comes in Section 5, where we estimate the error that is made by plugging  $\tilde{\theta}_m$  into the equation for  $\theta_m$ . The proof of Theorem 1.1 appears at the end of that section.

## 2. The fractal vector field: construction and regularity

In this section, we construct the periodic, incompressible vector field  $\mathbf{b}(\cdot)$  in Theorem 1.1 and prove that it is Hölder continuous.

### 2.1 A list of the ingredients used in the construction

We present here a list of the objects used in our construction of the incompressible vector field  $\mathbf{b}$ . These parameters are fixed throughout the rest of the paper.

- We let  $\beta$  be any positive exponent satisfying

$$1 < \beta < \frac{4}{3}. \quad (2.1)$$

This represents the regularity of the stream function  $\phi$  obtained in the construction. We typically think of  $\beta$  as very slightly smaller than  $4/3$ , perhaps 1.332. The parameter  $\alpha$  in the statement of Theorem 1.1 will be  $\alpha = \beta - 1$ .

- We define an exponent  $q$  explicitly in terms of  $\beta$  by

$$q := \frac{1}{2} \left( 1 + \frac{2 - \beta}{2(\beta - 1)} \right), \quad (2.2)$$

which prescribes the rate at which the scale separation  $\varepsilon_{m-1}/\varepsilon_m$  between successive scales  $\varepsilon_{m-1}$  and  $\varepsilon_m$  becomes larger as  $m$  becomes larger (and the scale  $\varepsilon_m$  becomes smaller): see (2.10), below. Equivalently, this means that

$$\beta = \frac{4}{3} \left( 1 - \frac{q - 1}{4q - 1} \right). \quad (2.3)$$

The main point is that  $q$  satisfies

$$1 < q < \frac{2 - \beta}{2(\beta - 1)}, \quad (2.4)$$

which follows from the inequality  $\beta < 4/3$  in (2.1).

- We fix the small parameter  $\delta \in (0, 1/16]$  defined explicitly by

$$\delta := \frac{1}{4} (q - 1) \left( 1 - \frac{2q + 1}{2q + 2} \beta \right) = \frac{(q - 1)^2}{4(q + 1)(4q - 1)}. \quad (2.5)$$

- We select a large positive integer  $N_* \in \mathbb{N}$  defined by

$$N_* := \left\lceil \frac{1}{\delta^2} + \frac{500}{\delta} \right\rceil. \quad (2.6)$$

The integer  $N_*$  counts the highest number of derivatives we need to track in our argument.

- We also define the special exponent

$$\gamma := \frac{(q - 1)\beta}{q + 1}. \quad (2.7)$$

This exponent is not used in the construction of the vector field in the next subsection, but it appears in Lemma 3.4 as a correction to the exponent for the renormalized diffusivities and subsequently in many of the computations in Section 4.

- We let  $\Lambda \in \mathbb{N} \cap [2^7, \infty)$  be a constant to be chosen later, at the very end of the arguments in Section 5. It is called the *minimal scale separation* and will be chosen to depend only on  $\beta$ .
- We define a sequences of length scales  $\{\varepsilon_m\}_{m \in \mathbb{N}_0}$  which satisfies

$$\varepsilon_m \ll \varepsilon_{m-1}^{1+\delta} \ll \varepsilon_{m-1},$$

in the sense that the separation between these length scales is at least a negative power of  $\varepsilon_m$ . These are defined as follows. We set  $\varepsilon_0 := 1$  and

$$\varepsilon_m^{-1} := \left\lceil \Lambda^{\frac{q^m}{q-1}} \right\rceil = \left\lceil \exp\left(\frac{q^m}{q-1} \log \Lambda\right) \right\rceil, \quad \forall m \in \mathbb{N}. \quad (2.8)$$

Since  $\varepsilon_1 \leq \Lambda^{-\frac{q}{q-1}} \leq \Lambda^{-1}$  and  $\left\lceil \Lambda^{\frac{q^{m+1}}{q-1}} \right\rceil = \left\lceil \Lambda^{\frac{q^m}{q-1}} \Lambda^{q^m} \right\rceil \geq \Lambda \left\lceil \Lambda^{\frac{q^m}{q-1}} \right\rceil$ , we have that

$$\frac{\varepsilon_m}{\varepsilon_{m+1}} \geq \Lambda, \quad \forall m \in \mathbb{N}_0. \quad (2.9)$$

In particular,  $\varepsilon_m \leq \Lambda^{-m} \leq 2^{-7m}$ . In fact, the sequence  $\{\varepsilon_m\}$  decreases at a super-geometric rate with  $\varepsilon_{m+1} \simeq \varepsilon_m^q$ , as it is routine to check that, for every  $m \geq 1$ ,

$$\frac{2}{3} \varepsilon_m^q \leq (1 - 10\varepsilon_m) \varepsilon_m^q \leq \varepsilon_{m+1} \leq (1 + 10\varepsilon_m) \varepsilon_m^q \leq \frac{4}{3} \varepsilon_m^q. \quad (2.10)$$

Finally, we remark that  $\varepsilon_m^{-1} \in \mathbb{N}$ .

- We define a sequence  $\{a_m\}_{m \in \mathbb{N}_0}$  of positive constants by

$$a_m := \varepsilon_m^{\beta-2}, \quad \forall m \in \mathbb{N}. \quad (2.11)$$

The constant  $a_m$  gives the strength of the shear flows at length scale  $\varepsilon_m$ : see (2.20), below.

- We introduce three sequences of time scales:  $\{\tau_m\}_{m \in \mathbb{N}_0}$ ,  $\{\tau'_m\}_{m \in \mathbb{N}}$  and  $\{\tau''_m\}_{m \in \mathbb{N}}$ . We define these in such a way that, for every  $m \in \mathbb{N}$ ,

$$\tau_m \ll \tau'_m \ll \tau''_m \ll a_{m-1}^{-1} \ll \tau_{m-1},$$

in the sense that the separation of the time scales in each of these inequalities is at least a negative power of  $\varepsilon_m$ . They are defined as follows: we set  $\tau_0 := 1$  and, for every  $m \in \mathbb{N}$ ,

$$\tau_m := \left(4 \left\lceil \varepsilon_{m-1}^{-\delta} \right\rceil + 1\right)^{-2} \tau''_m \quad (2.12)$$

$$\tau'_m := \left(4 \left\lceil \varepsilon_{m-1}^{-\delta} \right\rceil + 1\right)^{-1} \tau''_m, \quad (2.13)$$

$$\tau''_m := 2^{-25} \left\lceil \frac{a_{m-1}}{\varepsilon_{m-1}^{2\delta}} \right\rceil^{-1}. \quad (2.14)$$

In particular, we notice that

$$\frac{1}{\tau_m}, \frac{1}{\tau'_m}, \frac{1}{\tau''_m} \in 4\mathbb{N} \quad \text{and} \quad \frac{\tau'_m}{\tau_m}, \frac{\tau''_m}{\tau_m} \in 4\mathbb{N} + 1, \quad \forall m \in \mathbb{N}, \quad (2.15)$$

and

$$2^{-33} \varepsilon_{m-1}^{2-\beta+4\delta} \leq \tau_m \leq 2^{-28} \varepsilon_{m-1}^{2-\beta+4\delta} \quad \text{and} \quad 2^{-25} \varepsilon_{m-1}^{2-\beta+2\delta} \leq \tau''_m \leq 2^{-24} \varepsilon_{m-1}^{2-\beta+2\delta}. \quad (2.16)$$

Thus these sequences, like  $\{\varepsilon_m\}$ , are decreasing at a super-geometric rate. We also define

$$l_k := \left\lceil \frac{k\tau_m + \frac{1}{2}(\tau_m'' - \tau_m)}{\tau_m''} \right\rceil, \quad k \in \mathbb{Z}, \quad (2.17)$$

so that

$$k\tau_m + \left[-\frac{1}{2}\tau_m, \frac{1}{2}\tau_m\right] \subseteq l_k\tau_m'' + \left[-\frac{1}{2}\tau_m'', \frac{1}{2}\tau_m''\right], \quad \forall k \in \mathbb{Z}. \quad (2.18)$$

Notice that there are  $\frac{\tau_m''}{\tau_m}$  consecutive values of  $k$  that correspond to the same  $l_k$ .

- We define stream functions  $\psi_k$  for each  $k \in \mathbb{Z}$  by

$$\psi_{0,k}(x) := \begin{cases} \sin(2\pi x_i) & \text{if } k \in (4\mathbb{Z} + 2i - 1), i \in \{1, 2\}, \\ 0 & \text{if } k \in 2\mathbb{Z}. \end{cases} \quad (2.19)$$

The function  $\psi_{0,k}$  vanishes for even  $k$  and encodes a vertical shear for  $k \in 4\mathbb{Z} + 1$  and a horizontal shear for  $k \in 4\mathbb{Z} + 3$ . We scale these stream functions by defining

$$\psi_{m,k} := a_m \varepsilon_m^2 \psi_{0,k}\left(\frac{\cdot}{\varepsilon_m}\right), \quad \forall m \in \mathbb{N}, k \in \mathbb{Z}. \quad (2.20)$$

Note that the Lipschitz constant of the shear flows  $\nabla^\perp \psi_{m,k}$  is proportional to  $a_m$ . In fact, recalling that  $\llbracket \cdot \rrbracket_R$  defined in (1.21), we note that the stream functions  $\psi_{m,k}$  satisfy

$$\llbracket \psi_{m,k} \rrbracket_{\frac{2\pi}{\varepsilon_m}} \leq 5a_m \varepsilon_m^2, \quad \forall m \in \mathbb{N}, k \in \mathbb{Z}. \quad (2.21)$$

- We select an even cutoff function  $\zeta \in C_c^\infty(\mathbb{R})$  of time satisfying, for some constant  $C \in [1, \infty)$  which depends only on  $N_*$  (and thus only on  $\beta$ ),

$$0 \leq \zeta \leq \mathbf{1}_{[-\frac{2}{3}, \frac{2}{3}]}, \quad \sum_{k \in \mathbb{Z}} \zeta(\cdot - k) \equiv 1, \quad \max_{j \in \{0, \dots, N_*\}} \|\partial_t^j \zeta\|_{L^\infty(\mathbb{R})} \leq C, \quad (2.22)$$

and

$$\int_{\mathbb{R}} \zeta^2 = \frac{9}{10} \quad (2.23)$$

We also scale the cutoff function  $\zeta$  by setting

$$\zeta_{m,k}(t) := \zeta\left(\frac{t - k\tau_m}{\tau_m}\right), \quad \forall m \in \mathbb{N}, k \in \mathbb{Z}. \quad (2.24)$$

Observe that, for some constant  $C \in [1, \infty)$  depending only on  $\beta$  (see Remark 2.1 below),

$$0 \leq \zeta_{m,k} \leq \mathbf{1}_{[(k-\frac{2}{3})\tau_m, (k+\frac{2}{3})\tau_m]}, \quad \sum_{k \in \mathbb{Z}} \zeta_{m,k} = 1, \quad \max_{j \in \{0, \dots, N_*\}} \tau_m^j \|\partial_t^j \zeta_{m,k}\|_{L^\infty(\mathbb{R})} \leq C. \quad (2.25)$$

The role of these time cutoffs is to enable us to switch between the stream functions  $\psi_{m,k}$  for different  $k$ .

- We must define a *second* family of time cutoffs  $\{\xi_{m,k} : m \in \mathbb{N}_0, k \in \mathbb{Z}\}$  which live on slightly larger intervals than the  $\zeta_{m,k}$ 's. We take  $\xi \in C^\infty(\mathbb{R})$  to be a smooth cutoff function of time, satisfying, again for a constant  $C \in [1, \infty)$  depending only on  $\beta$ ,

$$\mathbf{1}_{[-\frac{3}{4}, \frac{3}{4}]} \leq \xi \leq \mathbf{1}_{[-\frac{5}{4}, \frac{5}{4}]}, \quad \sum_{k \in 2\mathbb{Z}+1} \xi(\cdot - k) \equiv 1 \quad \text{and} \quad \max_{j \in \{0, \dots, N_*\}} \|\partial_t^j \xi\|_{L^\infty(\mathbb{R})} \leq C \quad (2.26)$$

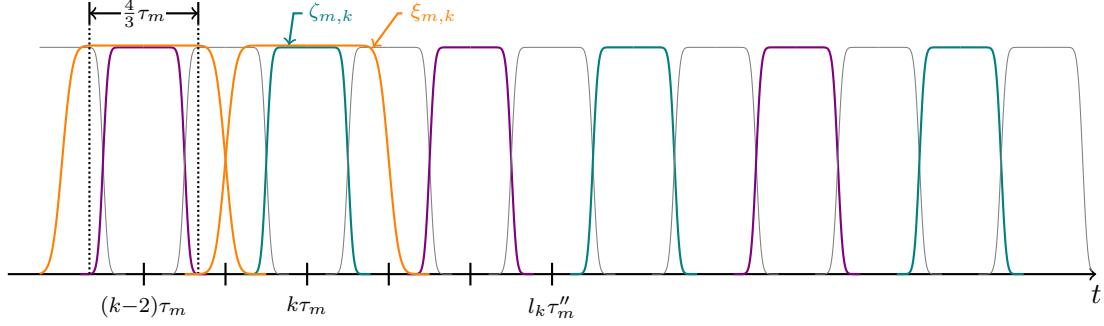


Figure 2.1: The families of cutoff functions  $\{\zeta_{m,k}\}$  and  $\{\xi_{m,k}\}$ . These are the small-scale time cutoff functions, and each of them is active on an interval of width close to  $\tau_m$ . The  $\zeta_{m,k}$ 's corresponding to odd  $k$  are drawn in colors—purple and green corresponding to horizontal and vertical shear flows, respectively—and in grey for even  $k$  as the corresponding vector fields vanish. Only two of the  $\xi_{m,k}$ 's are drawn (in orange). These have larger support than the  $\zeta_{m,k}$ 's and transition between zero and one on intervals in which the latter, for odd  $k$ , vanish.

and then define, for each  $k \in 4\mathbb{Z}$ ,

$$\xi_{m,k}(t) := \xi\left(\frac{t - k\tau_m}{\tau_m}\right). \quad (2.27)$$

Observe that the overlap in the periods between two successive  $\xi_{m,k}$ 's is disjoint from the support of the  $\zeta_{m,l}$ 's when  $l$  is odd (with some extra room). Precisely, we have

$$\text{dist}(\text{supp } \partial_t \xi_{m,k}, \text{supp } \zeta_{m,l}) \geq \frac{1}{12} \tau_m, \quad \forall k \in 2\mathbb{Z} + 1, \quad l \in 2\mathbb{Z} + 1. \quad (2.28)$$

See Figure 2.1 The cutoff functions  $\xi_{m,k}$  are not used in the construction of the vector field in the next subsection, but they are needed in the construction of the correctors in Section 3.1, see (3.12).

- We need to define two more families of time cutoffs

$$\{\hat{\zeta}_{m,l} : m \in \mathbb{N}_0, l \in \mathbb{Z}\} \quad \text{and} \quad \{\hat{\xi}_{m,l} : m \in \mathbb{N}_0, l \in \mathbb{Z}\}.$$

These have the following properties:

$$\hat{\zeta}_{m,l} = \hat{\zeta}_{m,0}(\cdot - l\tau''_m) \quad \text{and} \quad \hat{\xi}_{m,l} = \hat{\xi}_{m,0}(\cdot - l\tau''_m),$$

there exists  $C \in [1, \infty)$  depending only on  $\beta$  such that, for every  $m \in \mathbb{N}_0$  and  $l \in \mathbb{Z}$ ,

$$\begin{cases} \mathbf{1}_{[(l-\frac{1}{2})\tau''_m + 2\tau'_m, (l+\frac{1}{2})\tau''_m - 2\tau'_m]} \leq \hat{\zeta}_{m,l} \leq \mathbf{1}_{[(l-\frac{1}{2})\tau''_m + \tau'_m, (l+\frac{1}{2})\tau''_m - \tau'_m]}, \\ \mathbf{1}_{[(l-\frac{1}{2})\tau''_m + \tau'_m, (l+\frac{1}{2})\tau''_m - \tau'_m]} \leq \hat{\xi}_{m,l} \leq \mathbf{1}_{[(l-\frac{1}{2})\tau''_m - \tau'_m, (l+\frac{1}{2})\tau''_m + \tau'_m]}, \end{cases} \quad (2.29)$$

and

$$\sum_{l \in \mathbb{Z}} \hat{\xi}_{m,l} = 1, \quad (2.30)$$

and

$$\max_{j \in \{0, \dots, N_*\}} (\tau'_m)^j \left( \|\partial_t^j \hat{\zeta}_{m,l}\|_{L^\infty(\mathbb{R})} \vee \|\partial_t^j \hat{\xi}_{m,l}\|_{L^\infty(\mathbb{R})} \right) \leq C. \quad (2.31)$$

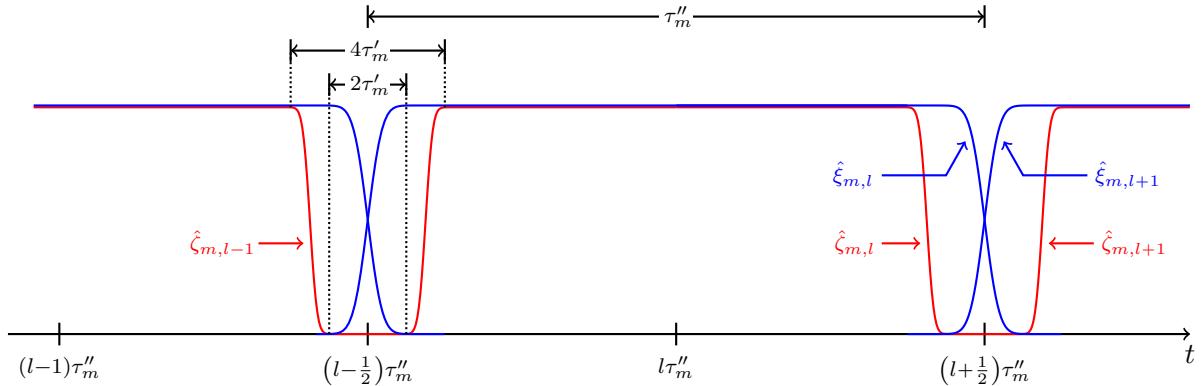


Figure 2.2: The families of cutoff functions  $\{\hat{\zeta}_{m,l}\}$  and  $\{\hat{\xi}_{m,l}\}$ . These are the large-scale time cutoff functions, and each of them is active on an interval of width close to  $\tau_m''$  and transition between zero and one in intervals of width  $\tau_m' \ll \tau_m''$ . The main difference is that the  $\hat{\xi}_{m,l}$ 's form a partition of unity, and their transition occurs entirely outside the support of the  $\hat{\zeta}_{m,l}$ 's.

Observe that  $\{\hat{\zeta}_{m,l}\}_{l \in \mathbb{Z}}$  does not form a partition of unity. In view of the definition (2.17), for every  $k, l \in \mathbb{Z}$ ,

$$l \neq l_k \implies \hat{\zeta}_{m,l} \zeta_{m,k} \equiv 0. \quad (2.32)$$

See Figure 2.2. The role of  $\hat{\zeta}_{m,l}$  is to smoothly cutoff the vector field near the time at which the flows and inverse flows are refreshed in our construction: see (2.36), below. The cutoff functions  $\hat{\xi}_{m,k}$  are not used in the construction of the vector field in the next subsection, but they are needed in the construction of the two-scale ansatz in Section 3.1, see (4.24).

**Remark 2.1** (Convention for the constants). Throughout the rest of the paper, we use  $C$  and  $c$  to denote positive constants which depend only on  $\beta$  and may vary in each occurrence. Note that, since the parameters  $q$ ,  $\delta$  and  $N_*$  are defined explicitly in terms of  $\beta$ , our constants may depend on them as well. In particular, these constants are understood to never depend on the scale parameter  $m$ . Also, since the parameter  $\Lambda$  will be chosen to be very large at the end of the proof of Theorem 1.1—precisely to absorb the error terms arising in the proof—it is important that we do not allow the constants  $C$  and  $c$  to depend on an upper bound for  $\Lambda$ . Since  $\Lambda \geq 32$ , they may depend on a lower bound for  $\Lambda$ .

## 2.2 Construction of the vector field

In this subsection, we construct an incompressible vector field  $\mathbf{b}$ , which is a sum of rescaled copies of a given family of periodic incompressible vector fields (shear flows) such that each term in the sum is advected by the partial sum of the terms representing larger scales (lower frequencies).

We proceed by constructing a sequence  $\{\mathbf{b}_m\}$  of smooth, incompressible vector fields which are  $\mathbb{Z} \times \mathbb{Z}^2$ -periodic, with associated sequences of periodic stream functions  $\{\phi_m\}$ , which satisfy

$$\nabla^\perp \phi_m = \mathbf{b}_m, \quad \langle \phi_m \rangle = 0, \quad (2.33)$$

and their corresponding flows  $\{X_m(\cdot, x, s)\}$  which satisfy, for each  $(x, s) \in \mathbb{R}^2 \times \mathbb{R}$ , the ODE

$$\begin{cases} \partial_t X_m(t, x, s) = \mathbf{b}_m(t, X_m(t, x, s)) & \text{in } \mathbb{R}, \\ X_m(s, x, s) = x. \end{cases} \quad (2.34)$$

We also denote by  $X_m^{-1}$  the inverse flow, that is,  $X_m^{-1}(t, \cdot, s)$  is the inverse function of  $X_m(t, \cdot, s)$ . The construction will be an iterative one, starting with the largest scale shears and progressively building in the smaller scale ones, in the Lagrangian coordinates of the larger scale shears. The vector field  $\mathbf{b}_m$  will have only scales larger than  $\varepsilon_m$  built into it, and we will eventually take  $\mathbf{b}$  as the limit  $m \rightarrow \infty$ .

We initialize the construction by setting

$$\begin{cases} \phi_0(t, x) := 0, \\ \mathbf{b}_0(t, x) := \nabla^\perp \phi_0 = 0, \\ X_0(t, x, s) := x. \end{cases} \quad (2.35)$$

Supposing that, for some  $m \in \mathbb{N}$ , we have defined  $\phi_j$ ,  $\mathbf{b}_j$  and  $X_j$  for every  $j \leq m-1$  and that these functions are smooth in all variables and satisfy the properties above. We then define  $\phi_m$ ,  $\mathbf{b}_m$  and  $X_m$  as follows. We first define the new stream function  $\phi_m$  by

$$\begin{aligned} \phi_m(t, x) &:= \phi_{m-1}(t, x) + \sum_{k, l \in \mathbb{Z}} \hat{\zeta}_{m, l}(t) \zeta_{m, k}(t) \psi_{m, k}(X_{m-1}^{-1}(t, x, l\tau_m'')) \\ &= \phi_{m-1}(t, x) + \sum_{k \in \mathbb{Z}} \hat{\zeta}_{m, l_k}(t) \zeta_{m, k}(t) \psi_{m, k}(X_{m-1}^{-1}(t, x, l_k\tau_m'')), \end{aligned} \quad (2.36)$$

where we recall that  $\psi_{m, k}$  is defined in (2.20). In the second equality in the display above we have appealed to (2.32). It is clear from induction that  $\phi_m$  is smooth. It also has zero mean by induction, the fact that  $\psi_{m, k}$  has zero mean and the fact that the incompressibility of  $\mathbf{b}_{m-1}$  implies that  $X_{m-1}^{-1}(t, \cdot, s)$  is measure-preserving. We may therefore define  $\mathbf{b}_m(t, x) := \nabla^\perp \phi_m(t, x)$  so that (2.33) is satisfied; and then define  $X_m(\cdot, x, s)$  to be the unique solution of the flow (2.34). These functions are clearly smooth, so this completes the construction.

Notice that the  $\psi_{m, k}$ 's change each time we increment  $k$ , but the inverse flows  $X_{m-1}^{-1}(\cdot, \cdot, l_k\tau_m'')$  depend on  $k$  only through the value of the initial time, namely  $l_k\tau_m''$ . In particular, the inverse flows are the same for  $\tau_m''/\tau_m \in 4\mathbb{N} + 1$  many consecutive values of  $k$ . To keep the notation short, we define, for every  $m \in \mathbb{N}$  and  $k \in \mathbb{Z}$ ,

$$X_{m, l}(t, x) := X_m(t, x, l\tau_m'') \quad \text{and} \quad X_{m, l}^{-1}(t, x) := X_m^{-1}(t, x, l\tau_m''). \quad (2.37)$$

As far as periodicity is concerned, it is clear from the construction that

$$\phi_m \quad \text{and} \quad \mathbf{b}_m \quad \text{are } \mathbb{Z} \times \mathbb{Z}^2\text{-periodic,}$$

that  $x \mapsto (X_m(t, x, s) - x)$  is  $\mathbb{Z}^2$ -periodic and  $X_m$  is 1-periodic in time, jointly in  $(t, s)$ , in the sense that

$$X_m(t + n, x, s + n) = X_m(t, x, s), \quad \forall m \in \mathbb{N}, n \in \mathbb{Z}, t, s \in \mathbb{R}, x \in \mathbb{R}^2.$$

We intend to define the vector field  $\mathbf{b}(t, x)$  by taking a limit:

$$\begin{cases} \phi(t, x) := \lim_{k \rightarrow \infty} \phi_k(t, x), \\ \mathbf{b}(t, x) := \nabla^\perp \phi(t, x). \end{cases} \quad (2.38)$$

The limit in the first line of (2.38) is valid in the sense of  $L^\infty([0, 1] \times [0, 1]^d)$  due to the fact that  $|\psi_{k, m}|$  is bounded by  $a_m \varepsilon_m^2 = \varepsilon_m^\beta$ , which is summable over  $m \in \mathbb{N}$ . That  $\phi$  is regular enough that the second line is valid is less clear at this stage.

### 2.3 Regularity of the stream functions and associated flows

How regular should we expect the stream function  $\phi$  to be? As we will discuss below, the regularity of  $\phi$  is complicated by the composition with the inverse flows in (2.36). It turns out that  $X_{m-1}^{-1}(t, \cdot, l_k \tau_m'')$  is close to the identity map on the support of the cutoff function  $\hat{\zeta}_{m,l_k}$ , as we will show below. If we imagine that the inverse flows can be replaced by the identity map in (2.36), then we may guess that the spatial regularity of  $\phi_m - \phi_{m-1}$  is similar to a periodic function with period  $\varepsilon_m$  and amplitude  $a_m \varepsilon_m^2 = \varepsilon_m^\beta$ . Thus, the  $C^{1,\beta'}$  seminorm of  $\phi_m - \phi_{m-1}$  should be of order  $\varepsilon_m^{\beta-\beta'-1}$ , for every  $\beta' < \beta - 1$ . Summing over the scales, this leads us to guess that  $\phi_m$  is uniformly bounded in  $C^{1,\beta'}$  for every  $\beta' < \beta - 1$ , and thus the limit  $\phi$  should belong to  $C^{1,\beta-}$ . This argument also suggests that  $\mathbf{b}$  should belong to  $C^{0,\beta-}$ .

This guess is correct, although the proof is more subtle than the back-of-the-envelope computation above may lead one to believe. Indeed, let us suppose that after the  $m$ th step of the construction we have  $\phi_m(t, \cdot) \in C^k$ , uniformly in  $t$  (with some estimate depending on  $m$  and  $k$ ). If we try to propagate this bound forward to  $\phi_{m+1}$ , what we find is that

$$\begin{aligned} \phi_m(t, \cdot) \in C^k &\implies \mathbf{b}_m(t, \cdot) \in C^{k-1} && \text{(by (2.33))}, \\ &\implies x \mapsto X_m^{-1}(t, x, s) \text{ is } C^{k-1} && \text{(regularity of transport equation)}, \\ &\implies \phi_{m+1} \in C^{k-1} && \text{(by the first line of (2.36))}. \end{aligned}$$

We have lost one derivative! This suggests that obtaining the desired uniform bounds on  $\phi_m$  requires propagating bounds on all spatial derivatives of  $\phi_m$ , using the analyticity of  $\psi_{k,m}$  (and its small size for large  $m$ ) to close the argument. This is the idea of the argument in the proof of Proposition 2.2, below. For this to succeed, every implication in the display above must be carefully quantified. Since the last implication uses the chain rule and must be iterated many times, we need a version of the Faá di Bruno formula (Proposition B.2) and the estimates for derivatives of the composition of two smooth functions that it implies (Proposition B.6). The second implication involves the regularity of the inverse flow  $X_{m-1}^{-1}$ , which solves the transport equation. We need estimates on all the derivatives of  $X_{m-1}^{-1}$  which are explicit in their dependence on the order of differentiation. Such estimates are classical, but also difficult to find in the literature in the explicit form we require here; for the reader's convenience, in Appendix B we present complete statements and proofs of what we need here.

We recall that by (2.16) and (2.29), if  $t \in \text{supp } \hat{\zeta}_{m,l_k}$ , then  $|t - l_k \tau_m''| \leq \frac{1}{2} \tau_m'' \leq 2^{-25} \varepsilon_m^{2-\beta+2\delta} \leq 2^{-25} \varepsilon_m^{2-\beta}$ . As such each flow  $X_m(t+s, x, s)$  and backwards flow  $X_m^{-1}(t+s, x, s)$  needs to be studied for times  $t$  which satisfy  $|t| \leq 2^{-25} \varepsilon_m^{2-\beta}$ . This motivates assumption (2.40) below. Also, recall that  $\llbracket \cdot \rrbracket_{n,R}$  is defined in (1.21).

**Proposition 2.2.** *For every  $m \in \mathbb{N}$ ,*

$$\sup_t \llbracket (\phi_m - \phi_{m-1})(\cdot, t) \rrbracket_{n, 2^7 \varepsilon_m^{-1}} \leq 10 \varepsilon_m^\beta, \quad \forall n \in \mathbb{N}_0, \quad (2.39)$$

and, for every  $s, t \in \mathbb{R}$  with  $|t| \leq 2^{-25} a_m^{-1} = 2^{-25} \varepsilon_m^{2-\beta}$ ,

$$\llbracket \nabla X_m^{-1}(t+s, \cdot, s) - \mathbf{I}_2 \rrbracket_{n, 2^{11} \varepsilon_m^{-1}} \leq 2^{23} |t| a_m = 2^{23} |t| \varepsilon_m^{\beta-2}, \quad \forall n \in \mathbb{N}_0. \quad (2.40)$$

Consequently, for any  $\beta' \in (0, \beta - 1]$  we have

$$\|\phi_m - \phi_{m-1}\|_{L^\infty(\mathbb{R}; C^{1,\beta'}(\mathbb{R}^2))} \leq 2^{12} \varepsilon_m^{\beta-1-\beta'}, \quad (2.41)$$

and there exists a constant  $C > 0$  which only depends on  $\beta$ , such that

$$\|\phi_m - \phi_{m-1}\|_{C^{1,\beta'}(\mathbb{R}; L^\infty(\mathbb{R} \times \mathbb{R}^2))} \leq C \varepsilon_m^{\beta-1-\beta'}. \quad (2.42)$$

*Proof of Proposition 2.2.* Let  $\{M_m\}_{m \in \mathbb{N}} \subseteq (0, \infty)$  and  $\{R_m\}_{m \in \mathbb{N}} \subseteq (0, \infty)$  be two sequences to be defined explicitly below (see (2.49)). Suppose that, for some  $m \in \mathbb{N}$ , we have

$$\tau_m'' \leq (32M_{m-1})^{-1} \quad (2.43)$$

and

$$\sup_{t \in \mathbb{R}} \llbracket \phi_{m-1} \rrbracket_{n, R_{m-1}} \leq M_{m-1} R_{m-1}^{-2} \frac{(n+2)^2}{(n+1)^3}, \quad \forall n \in \mathbb{N}, n \geq 2. \quad (2.44)$$

The assumption (2.44) implies, for every  $n \in \mathbb{N}$ ,

$$\begin{aligned} \sup_{t \in \mathbb{R}} \llbracket \mathbf{b}_{m-1} \rrbracket_{n, R_{m-1}} &\leq \frac{(n+1)^2}{n! R_{m-1}^n} \sup_{t \in \mathbb{R}} \|\nabla^{n+1} \phi_{m-1}\|_{L^\infty(\mathbb{R}^2)} \\ &\leq \frac{(n+1)^2}{n! R_{m-1}^n} \sup_{t \in \mathbb{R}} \llbracket \phi_{m-1} \rrbracket_{n+1, R_{m-1}} \frac{(n+1)! R_{m-1}^{n+1}}{(n+2)^2} \\ &\leq M_{m-1} R_{m-1}^{-1}. \end{aligned} \quad (2.45)$$

According to the definition (2.34), the backwards flow  $X_{m-1}^{-1}$  solves  $(\partial_t + \mathbf{b}_{m-1} \cdot \nabla) X_{m-1}^{-1} = 0$ , and so with  $Y = X_{m-1}^{-1}(\cdot, x, \cdot) - x$ , we are in the setting of Lemma B.7, with  $\mathbf{f} = -\mathbf{g} = \mathbf{b}_{m-1}$ . The previously established estimate (2.45) shows that assumptions (B.12) hold with  $C_{\mathbf{f}} = C_{\mathbf{g}} = M_{m-1} R_{m-1}^{-1}$ , and  $R_{\mathbf{f}} = R_{\mathbf{g}} = R_{m-1}$ . Here we emphasize that the assumption (B.12) is only required to hold for derivative indices  $n$  with  $n \geq 1$ . According to (B.14), we thus define  $\bar{R}_{m-1}(t) := R_{m-1}(1 + 8|t|M_{m-1})$ . By Lemma B.7, for every  $s, t \in \mathbb{R}$  such that  $|t| \leq (8M_{m-1})^{-1}$ , and for every  $n \in \mathbb{N}$ ,

$$\llbracket X_{m-1}^{-1}(t+s, \cdot, s) - x \rrbracket_{n, \bar{R}_{m-1}(t)} \leq 16|t|M_{m-1} R_{m-1}^{-1}. \quad (2.46)$$

Observe that if  $|t| \leq (64M_{m-1})^{-1}$ , then  $\bar{R}_{m-1}(t) \leq \frac{9}{8}R_{m-1}$  and (2.46) implies that for all  $n \geq 1$

$$\begin{aligned} \llbracket X_{m-1}^{-1}(t+s, \cdot, s) \rrbracket_{n, \frac{9}{8}R_{m-1}} &\leq \llbracket X_{m-1}^{-1}(t+s, \cdot, s) - x \rrbracket_{n, \bar{R}_{m-1}(t)} + \llbracket x \rrbracket_{n, \frac{9}{8}R_{m-1}} \\ &\leq 16|t|M_{m-1} R_{m-1}^{-1} + 4 \cdot \frac{8}{9}R_{m-1}^{-1} \leq 5 \cdot \frac{8}{9}R_{m-1}^{-1}. \end{aligned} \quad (2.47)$$

We now have all the necessary ingredients for estimating the second term in (2.36) using Proposition B.6. With the help of (2.21), (2.43), (2.47), and (B.9) we obtain, for every  $t \in \mathbb{R}$  and every  $n \in \mathbb{N}$  with  $n \geq 2$  that

$$\begin{aligned} \llbracket (\phi_m - \phi_{m-1})(t, \cdot) \rrbracket_{n, \frac{9}{8}R_{m-1} + \frac{20\pi}{\varepsilon_m}} &\leq 2 \sup_{k \in \mathbb{Z}} \sup_{t \in \text{supp } \hat{\zeta}_{m, l_k} \zeta_{m, k}} \llbracket \psi_{m, k} \circ X_{m-1}^{-1}(t, \cdot, l_k \tau_m'') \rrbracket_{n, \frac{9}{8}R_{m-1} + \frac{20\pi}{\varepsilon_m}} \\ &\leq 10a_m \varepsilon_m^2. \end{aligned} \quad (2.48)$$

Here we have used that if  $t \in \text{supp } \hat{\zeta}_{m, l_k}$ , then by (2.29) we have that  $|t - l_k \tau_m''| \leq \frac{1}{2} \tau_m''$ , and thus (2.43) implies  $|t - l_k \tau_m''| \leq (64M_{m-1})^{-1}$ . Therefore, by the induction assumption (2.44), the bound (2.48), and the monotonicity of  $\llbracket \cdot \rrbracket_{n, R}$  with respect to  $R$ , for all  $n \geq 2$ ,  $t \in \mathbb{R}$ , and  $R_m \geq \frac{9}{8}R_{m-1} + \frac{20\pi}{\varepsilon_m} > R_{m-1}$ , we have

$$\begin{aligned} \llbracket \phi_m(t, \cdot) \rrbracket_{n, R_m} &\leq \llbracket \phi_{m-1}(t, \cdot) \rrbracket_{n, R_m} + \llbracket (\phi_m - \phi_{m-1})(t, \cdot) \rrbracket_{n, R_m} \\ &\leq M_{m-1} R_{m-1}^{-2} \frac{(n+2)^2}{(n+1)^3} \left( \frac{R_{m-1}}{R_m} \right)^n + 10a_m \varepsilon_m^2 \left( \frac{\frac{9}{8}R_{m-1} + \frac{20\pi}{\varepsilon_m}}{R_m} \right)^n \\ &\leq M_m R_m^{-2} \frac{(n+2)^2}{(n+1)^3} \left( \frac{M_{m-1}}{M_m} \left( \frac{R_{m-1}}{R_m} \right)^{n-2} + \frac{10a_m \varepsilon_m^2 R_m^2 (n+1)}{M_m} \left( \frac{\frac{9}{8}R_{m-1} + \frac{20\pi}{\varepsilon_m}}{R_m} \right)^n \right). \end{aligned}$$

Thus, if we make the choice  $R_m \geq 2(\frac{9}{8}R_{m-1} + \frac{20\pi}{\varepsilon_m})$ , since  $(n+1)2^{-n+2} \leq 3$  for  $n \geq 2$ , we arrive at

$$[\phi_m(t, \cdot)]_{n, R_m} \leq M_m R_m^{-2} \frac{(n+2)^2}{(n+1)^3} \left( \frac{M_{m-1}}{M_m} + \frac{30a_m \varepsilon_m^2 (\frac{9}{8}R_{m-1} + \frac{20\pi}{\varepsilon_m})^2}{M_m} \right).$$

With an eye on the induction hypothesis (2.44) with  $m$  in place of  $m-1$ , this motivates the recursion

$$\begin{cases} R_m := \frac{9}{4}R_{m-1} + 2^7 \varepsilon_m^{-1} \\ M_m := M_{m-1} + 2^7 a_m \varepsilon_m^2 R_{m-1}^2 + 2^{18} a_m, \end{cases} \quad (2.49)$$

for all  $m \geq 1$ . Here we have used that  $40\pi \leq 2^7$ ,  $60 \cdot (\frac{9}{8})^2 \leq 2^7$ , and  $60 \cdot (20\pi)^2 \leq 2^{18}$ . We now take the recurrence (2.49) to be the definition of the sequences  $\{M_m\}$  and  $\{R_m\}$ , starting from  $M_0 := 1$  and  $R_0 := 1$ .

We next analyze the recurrence relation (2.49). Observe that

$$\varepsilon_{m+1} R_m = \left( \frac{9\varepsilon_{m+1}}{4\varepsilon_m} \right) \varepsilon_m R_{m-1} + 2^7 \left( \frac{\varepsilon_{m+1}}{\varepsilon_m} \right).$$

Recall from (2.9) and the fact that  $\Lambda \geq 2^7$  that

$$\frac{\varepsilon_{m+1}}{\varepsilon_m} \leq \frac{1}{2^7}, \quad \forall m \in \mathbb{N}. \quad (2.50)$$

We deduce by induction that, for every  $m \in \mathbb{N}$  with  $m \geq 1$ ,

$$\varepsilon_{m+1} R_m \leq \frac{5}{4}. \quad (2.51)$$

Inserting the bound (2.51) back into the first line of (2.49) yields

$$R_m \leq (3 + 2^7) \varepsilon_m^{-1}. \quad (2.52)$$

Inserting the bound (2.51) into second line in (2.49), we obtain that

$$M_m \leq M_{m-1} + (2^8 + 2^{18}) a_m,$$

and thus by appealing to (2.9),  $\beta < 4/3$ ,  $\Lambda \geq 2^7$ , and (2.11), for every  $m \in \mathbb{N}$  we obtain the bound

$$M_m \leq (2^8 + 2^{18}) \sum_{j=0}^m a_j \leq (2^8 + 2^{18}) a_m \sum_{j \geq 0} \Lambda^{j(\beta-2)} \leq \frac{2^8 + 2^{18}}{1 - \Lambda^{\beta-2}} a_m \leq \frac{2^8 + 2^{18}}{1 - 2^{-14/3}} a_m \leq 2^{19} a_m. \quad (2.53)$$

In view of (2.16), using (2.52) and (2.53) we find that

$$\tau_m'' \leq 2^{-24} a_{m-1}^{-1} \varepsilon_{m-1}^\delta \leq 2^{-24} a_{m-1}^{-1} \leq 2^{-5} M_{m-1}^{-1}.$$

That is, the hypothesis (2.43) is in fact valid for every  $m \in \mathbb{N}$  and is therefore superfluous.

By induction, we may conclude now that (2.44) holds for every  $m \in \mathbb{N}$ . Moreover, we have shown that (2.47) and (2.48) are valid for every  $m \in \mathbb{N}$ , for  $n \geq 1$  and respectively  $n \geq 2$ .

Substituting (2.52), (2.53), and the bound  $2^7\varepsilon_m^{-1} \leq R_m \leq 2^8\varepsilon_m^{-1}$  into the these bounds yields, for  $m \in \mathbb{N}$  and all  $n \in \mathbb{N}$  with  $n \geq 2$ , that

$$\sup_{t \in \mathbb{R}} \llbracket \phi_m(t, \cdot) \rrbracket_{n, 2^8\varepsilon_m^{-1}} \leq 2^5 a_m \varepsilon_m^2 \frac{(n+2)^2}{(n+1)^3}, \quad (2.54)$$

$$\sup_{t \in \mathbb{R}} \llbracket (\phi_m - \phi_{m-1})(t, \cdot) \rrbracket_{n, 2^7\varepsilon_m^{-1}} \leq 10 a_m \varepsilon_m^2 = 10 \varepsilon_m^\beta, \quad (2.55)$$

and for every  $s, t \in \mathbb{R}$  with  $|t| \leq (2^{25}a_m)^{-1}$ , and all  $n \in \mathbb{N}$ , that

$$\llbracket X_m^{-1}(t+s, \cdot, s) - x \rrbracket_{n, 2^9\varepsilon_m^{-1}} \leq 2^{16} |t| a_m \varepsilon_m. \quad (2.56)$$

The bound (2.55) implies (2.39) for  $n \in \mathbb{N}$  with  $n \geq 2$ , while the estimate (2.56) yields (2.40) for every  $n \in \mathbb{N}_0$ , upon noting that

$$\llbracket \nabla X_m^{-1}(t+s, \cdot, s) - I_2 \rrbracket_{n, 2^{11}\varepsilon_m^{-1}} \leq \frac{(n+1)^{32^{9-2n}\varepsilon_m^{-1}}}{(n+2)^2} \llbracket X_m^{-1}(t+s, \cdot, s) - \cdot \rrbracket_{n+1, 2^9\varepsilon_m^{-1}} \leq 2^{23} |t| a_m.$$

In order to get (2.39) for  $n = 0$ , we use the definition of  $\phi_m$  in (2.36) and obtain

$$\|\phi_m - \phi_{m-1}\|_{L^\infty(\mathbb{R} \times \mathbb{R}^2)} \leq a_m \varepsilon_m^2 = \varepsilon_m^\beta.$$

For the  $n = 1$  bound, we interpolate between the above estimate and (2.55) with  $n = 2$ , to obtain

$$\|\nabla(\phi_m - \phi_{m-1})\|_{L^\infty(\mathbb{R} \times \mathbb{R}^2)} \leq (\varepsilon_m^\beta)^{1/2} (10 a_m \varepsilon_m^2 \frac{2}{9} \cdot (2^7 \varepsilon_m^{-1})^2)^{1/2} \leq 2^8 \varepsilon_m^{\beta-1}, \quad (2.57)$$

and thus

$$\sup_{t \in \mathbb{R}} \llbracket (\phi_m - \phi_{m-1})(t, \cdot) \rrbracket_{1, 2^7\varepsilon_m^{-1}} \leq 8 \varepsilon_m^\beta \leq 10 \varepsilon_m^\beta.$$

Thus we have also proved (2.39) for  $n \in \{0, 1\}$ , and hence in view of (2.55) for every  $n \in \mathbb{N}_0$ .

For future purposes, we note at this stage that upon telescoping the bound (2.39) for  $n \in \{0, 1\}$ , similarly to (2.53) we obtain

$$\begin{aligned} \sup_{t \in \mathbb{R}} \llbracket \phi_m(t, \cdot) \rrbracket_{n, 2^8\varepsilon_m^{-1}} &\leq \sum_{j=0}^m \sup_{t \in \mathbb{R}} \llbracket (\phi_j - \phi_{j-1})(t, \cdot) \rrbracket_{n, 2^7\varepsilon_j^{-1}} \left( \frac{\varepsilon_j^{-1}}{2\varepsilon_m^{-1}} \right)^n \\ &\leq 10 \cdot 2^{-n} \varepsilon_m^n \sum_{j=0}^m \varepsilon_j^{\beta-n} \leq 10 \cdot 2^{-n} \varepsilon_m^n \sum_{j=0}^m \Lambda^{-j(\beta-n)} \leq \begin{cases} 11, & n = 0, \\ \frac{3}{\beta-1} \varepsilon_m, & n = 1, \end{cases} \end{aligned} \quad (2.58)$$

for  $n \in \{0, 1\}$ . Here we have used that  $\beta > 1$  and  $\Lambda \geq 2^7$ . The above estimate with  $n = 1$  then immediately implies

$$\|\mathbf{b}_m\|_{L^\infty(\mathbb{R} \times \mathbb{R}^2)} \leq 2^7 \varepsilon_m^{-1} \sup_t \llbracket \phi_m \rrbracket_{1, 2^8\varepsilon_m^{-1}} \leq 2^{9/(\beta-1)}, \quad (2.59)$$

which shows that the sequence of vector fields  $\{\mathbf{b}_m\}_{m \geq 0}$  is uniformly bounded in space-time, uniformly in  $m$ .

In order to conclude the proof, we need to still consider the bounds (2.41) and (2.42). The first one, namely (2.41), follows by interpolating the bounds in (2.39) when  $n = 1$  and  $n = 2$ . For the Hölder regularity of the time derivative, we differentiate the expression (2.36) in time to obtain

$$\begin{aligned} \partial_t(\phi_m - \phi_{m-1}) &= \sum_{k \in \mathbb{Z}} \partial_t(\hat{\zeta}_{m, l_k} \zeta_{m, k}) \psi_{m, k} \circ X_{m-1, l_k}^{-1} \\ &\quad - \sum_{k \in \mathbb{Z}} \hat{\zeta}_{m, l_k} \zeta_{m, k} \nabla \psi_{m, k} \circ X_{m-1, l_k}^{-1} \cdot \mathbf{b}_{m-1} \cdot \nabla X_{m-1, l_k}^{-1}. \end{aligned} \quad (2.60)$$

Using (2.20), (2.25), (2.31), (2.59), and (2.57), we obtain

$$\begin{aligned}
& \|\partial_t(\phi_m - \phi_{m-1})\|_{L^\infty(\mathbb{R} \times \mathbb{R}^2)} \\
& \leq 2\|\partial_t(\hat{\zeta}_{m,l_k} \zeta_{m,k})\|_{L^\infty(\mathbb{R})} \sup_{k \in \mathbb{Z}} \|\psi_{m,k}\|_{L^\infty(\mathbb{R}^2)} \\
& \quad + 2\|\mathbf{b}_{m-1}\|_{L^\infty(\mathbb{R} \times \mathbb{R}^2)} \sup_{k \in \mathbb{Z}} \|\nabla \psi_{m,k}\|_{L^\infty(\mathbb{R}^2)} \sup_{t \in \text{supp } \hat{\zeta}_{m,l_k} \zeta_{m,k}} \|\nabla X_{m-1,l_k}^{-1}\|_{L^\infty(\mathbb{R}^2)} \\
& \leq C a_m \varepsilon_m^2 \tau_m^{-1} + C a_m \varepsilon_m \\
& \leq C(\varepsilon_{m-1}^{q-(2-\beta+4\delta)} + 1) \varepsilon_m^{\beta-1} \leq C \varepsilon_m^{\beta-1}. \tag{2.61}
\end{aligned}$$

The exponent of  $\varepsilon_{m-1}$  in the last line was computed using (2.3), (2.5), and (2.10). Here we have also used that  $\beta > 1$  and  $q > 1$ . By applying  $\nabla^\perp$  to (2.60), similarly to (2.61) we deduce

$$\begin{aligned}
& \|\partial_t(\mathbf{b}_m - \mathbf{b}_{m-1})\|_{L^\infty(\mathbb{R} \times \mathbb{R}^2)} \\
& \leq 2\|\partial_t(\hat{\zeta}_{m,l_k} \zeta_{m,k})\|_{L^\infty(\mathbb{R})} \sup_{k \in \mathbb{Z}} \|\nabla \psi_{m,k}\|_{L^\infty(\mathbb{R}^2)} \sup_{t \in \text{supp } \hat{\zeta}_{m,l_k} \zeta_{m,k}} \|\nabla X_{m-1,l_k}^{-1}\|_{L^\infty(\mathbb{R}^2)} \\
& \quad + 2\|\mathbf{b}_{m-1}\|_{L^\infty(\mathbb{R} \times \mathbb{R}^2)} \sup_{k \in \mathbb{Z}} \|\nabla^2 \psi_{m,k}\|_{L^\infty(\mathbb{R}^2)} \sup_{t \in \text{supp } \hat{\zeta}_{m,l_k} \zeta_{m,k}} \|\nabla X_{m-1,l_k}^{-1}\|_{L^\infty(\mathbb{R}^2)}^2 \\
& \quad + 2\|\mathbf{b}_{m-1}\|_{L^\infty(\mathbb{R} \times \mathbb{R}^2)} \sup_{k \in \mathbb{Z}} \|\nabla \psi_{m,k}\|_{L^\infty(\mathbb{R}^2)} \sup_{t \in \text{supp } \hat{\zeta}_{m,l_k} \zeta_{m,k}} \|\nabla^2 X_{m-1,l_k}^{-1}\|_{L^\infty(\mathbb{R}^2)} \\
& \quad + 2\|\nabla \mathbf{b}_{m-1}\|_{L^\infty(\mathbb{R} \times \mathbb{R}^2)} \sup_{k \in \mathbb{Z}} \|\nabla \psi_{m,k}\|_{L^\infty(\mathbb{R}^2)} \sup_{t \in \text{supp } \hat{\zeta}_{m,l_k} \zeta_{m,k}} \|\nabla X_{m-1,l_k}^{-1}\|_{L^\infty(\mathbb{R}^2)} \\
& \leq C a_m \varepsilon_m \tau_m^{-1} + C a_m + C a_m \varepsilon_m a_{m-1} \\
& \leq C a_m = C \varepsilon_m^{\beta-2}.
\end{aligned}$$

The upshot of the above estimate is that, after telescoping, we arrive at

$$\|\partial_t \mathbf{b}_m\|_{L^\infty(\mathbb{R} \times \mathbb{R}^2)} \leq C \sum_{j=0}^m \varepsilon_j^{\beta-2} \leq C \varepsilon_m^{\beta-2}. \tag{2.62}$$

Next, we apply one more time derivative to (2.60) to obtain

$$\begin{aligned}
& \partial_t^2(\phi_m - \phi_{m-1})(t, x) \\
& = \sum_{k \in \mathbb{Z}} \partial_t^2(\hat{\zeta}_{m,l_k} \zeta_{m,k}) \psi_{m,k} \circ X_{m-1,l_k}^{-1} \\
& \quad - 2 \sum_{k \in \mathbb{Z}} \partial_t(\hat{\zeta}_{m,l_k} \zeta_{m,k}) \nabla \psi_{m,k} \circ X_{m-1,l_k}^{-1} \cdot \mathbf{b}_{m-1} \cdot \nabla X_{m-1,l_k}^{-1} \\
& \quad + \sum_{k \in \mathbb{Z}} \hat{\zeta}_{m,l_k} \zeta_{m,k} \nabla^2 \psi_{m,k} \circ X_{m-1,l_k}^{-1} : (\mathbf{b}_{m-1} \cdot \nabla X_{m-1,l_k}^{-1}) \otimes (\mathbf{b}_{m-1} \cdot \nabla X_{m-1,l_k}^{-1}) \\
& \quad - \sum_{k \in \mathbb{Z}} \hat{\zeta}_{m,l_k} \zeta_{m,k} \nabla \psi_{m,k} \circ X_{m-1,l_k}^{-1} \cdot \left( \partial_t \mathbf{b}_{m-1} \cdot \nabla X_{m-1,l_k}^{-1} - \mathbf{b}_{m-1} \cdot \nabla(\mathbf{b}_{m-1} \cdot \nabla X_{m-1,l_k}^{-1}) \right). \tag{2.63}
\end{aligned}$$

Similarly to (2.61), and appealing in addition to the estimate (2.62), we obtain

$$\begin{aligned}
\|\partial_t^2(\phi_m - \phi_{m-1})\|_{L^\infty(\mathbb{R} \times \mathbb{R}^2)} & \leq C a_m \varepsilon_m^2 \tau_m^{-2} + C a_m \varepsilon_m \tau_m^{-1} + C a_m + C a_m \varepsilon_m (\varepsilon_{m-1}^{\beta-2} + a_{m-1}) \\
& \leq C a_m = C \varepsilon_m^{\beta-2}. \tag{2.64}
\end{aligned}$$

The claimed estimate (2.42) now follows from (2.61) and (2.64) by interpolation, concluding the proof of the Proposition.  $\square$

**Corollary 2.3.** *There exists a  $C \in [1, \infty)$  which only depends on  $\beta$ , such that for every  $m \in \mathbb{N}$ ,*

$$\sup_{t \in \mathbb{R}} \llbracket \phi_m(\cdot, t) \rrbracket_{n, C\varepsilon_m^{-1}} \leq C\varepsilon_m^\beta, \quad \forall n \in \mathbb{N}, n \geq 2, \quad (2.65)$$

and

$$\sup_{t \in \mathbb{R}} \llbracket \phi_m(\cdot, t) \rrbracket_{n, C\varepsilon_m^{-1}} \leq C\varepsilon_m^n, \quad \forall n \in \{0, 1\}. \quad (2.66)$$

For every  $\beta' \in (0, \beta - 1)$ , the stream function  $\phi$  belongs to  $C^0(\mathbb{R}; C^{1, \beta'}(\mathbb{R}^2)) \cap C^{1, \beta'}(\mathbb{R}; L^\infty(\mathbb{R}^2))$ , and in particular, the vector field  $\mathbf{b}$  belongs to  $C^0(\mathbb{R}; C^{0, \beta'}(\mathbb{R}^2)) \cap C^{0, \beta'}(\mathbb{R}; L^\infty(\mathbb{R}^2))$ .

*Proof of Corollary 2.3.* The bounds for the derivatives of  $\phi_m$  of order  $n$  with  $n \geq 2$ , claimed in (2.65), were already established in (2.54). The estimate (2.66) was proven earlier in (2.58). The claimed regularity of  $\phi = \lim_{m \rightarrow \infty} \phi_m$  follows by telescoping sum  $\phi_m = \sum_{j=1}^m (\phi_j - \phi_{j-1})$ , appealing to (2.41) and (2.42), and using the fact that by (2.9) we have

$$\sum_{n=1}^m \varepsilon_m^n \leq C|p|^{-1} \cdot \begin{cases} 1 & \text{if } p > 0, \\ \varepsilon_m^p & \text{if } p < 0. \end{cases}$$

The regularity of  $\mathbf{b} = \nabla^\perp \phi = \lim_{m \rightarrow \infty} \mathbf{b}_m$  follows from that of  $\phi$  by interpolation.  $\square$

By combining the estimates established in Proposition 2.2 with the results of Proposition B.10, we obtain the following useful results.

**Corollary 2.4.** *For every  $s, t \in \mathbb{R}$  with  $|t| \leq 2^{-25}a_m^{-1}$ , we have that*

$$\|\nabla X_m(t + s, \cdot, s) - \mathbf{I}_2\|_{L^\infty(\mathbb{R}^2)} \leq 2^{23}|t|a_m \leq \frac{1}{4}. \quad (2.67)$$

Moreover, for all  $s, t \in \mathbb{R}$  with  $|t| \leq 2^{-25}a_m^{-1}$ , we have

$$\llbracket \nabla X_m(t + s, X_m^{-1}(t + s, \cdot, s), s) - \mathbf{I}_2 \rrbracket_{n, 2^{10}\varepsilon_m^{-1}} \leq 40, \quad \forall n \in \mathbb{N}_0, \quad (2.68)$$

and

$$\llbracket \nabla X_m(t + s, \cdot, s) \rrbracket_{n, 2^{14}\varepsilon_m^{-1}} \leq 12, \quad \forall n \in \mathbb{N}. \quad (2.69)$$

In particular, (2.67) and (2.69) imply that

$$\sup_{|t| \leq 2^{-25}a_m^{-1}} \|\nabla^n X_m(t + s, \cdot, s)\|_{L^\infty(\mathbb{R}^2)} \leq 2n!(2^{13}\varepsilon_m^{-1})^{n-1}, \quad \forall n \in \mathbb{N}. \quad (2.70)$$

*Proof of Corollary 2.4.* From (2.54), we deduce that for any  $n \in \mathbb{N}$ ,

$$\begin{aligned} \sup_{t \in \mathbb{R}} \llbracket \mathbf{b}_m(t, \cdot) \rrbracket_{n, 2^8\varepsilon_m^{-1}} &\leq \sup_{t \in \mathbb{R}} \frac{2(n+1)^2}{n!(2^8\varepsilon_m^{-1})^n} \|\nabla^{n+1} \phi_m(t, \cdot)\|_{L^\infty(\mathbb{R}^2)} \\ &\leq \sup_{t \in \mathbb{R}} \llbracket \phi_m(t, \cdot) \rrbracket_{n+1, 2^8\varepsilon_m^{-1}} \frac{2(n+1)^2}{n!(2^8\varepsilon_m^{-1})^n} \frac{(n+1)!(2^8\varepsilon_m^{-1})^{n+1}}{(n+2)^2} \\ &\leq 2^5 a_m \varepsilon_m^2 \frac{(n+2)^2}{(n+1)^3} \frac{2(n+1)^3(2^8\varepsilon_m^{-1})}{(n+2)^2} \leq 2^{14} a_m \varepsilon_m. \end{aligned} \quad (2.71)$$

The definition (2.34) suggests that we apply Proposition B.10 with  $\mathbf{f} = \mathbf{b}_m$ ,  $C_{\mathbf{f}} = 2^{14}a_m\varepsilon_m$ , and  $R_{\mathbf{f}} = 2^8\varepsilon_m^{-1}$ . The bound (B.25) then directly implies (2.67) since  $|t| \leq 2^{-25}a_m^{-1} = (8C_{\mathbf{f}}R_{\mathbf{f}})^{-1}$ . Similarly, the bound (2.68) for  $n \geq 1$  follows from (B.26) with  $d = 2$ , since  $R_{\mathbf{f}}(1 + 8|t|C_{\mathbf{f}}R_{\mathbf{f}})^2 \leq 4R_{\mathbf{f}} = 2^{10}\varepsilon_m^{-1}$ . The bound (2.69) is a direct consequence of (B.28) for  $d = 2$ , since  $16R_{\mathbf{f}}(1 + 16C_{\mathbf{f}}R_{\mathbf{f}}|t|) \leq 48R_{\mathbf{f}} \leq 2^{14}\varepsilon_m^{-1}$ . Lastly, the estimate (2.70) follows from (2.67) (for  $n = 1$ ) and (2.69) (for  $n \geq 2$ ), upon recalling definition (1.21).  $\square$

## 2.4 Material derivative estimates

The bound (2.71) implies that for all  $n \in \mathbb{N}$  we have

$$\|\nabla^n \mathbf{b}_m\|_{L^\infty(\mathbb{R} \times \mathbb{R}^2)} \leq 2^{14} a_m \varepsilon_m \frac{n!(2^8 \varepsilon_m^{-1})^n}{(n+1)^2} \leq 2^{22} \varepsilon_m^{\beta-2} (n-1)! (2^8 \varepsilon_m^{-1})^{n-1}. \quad (2.72)$$

In order to estimate the time correctors in our two-scale ansatz in Section 4, it turns out that we also need to have estimates available for

$$\|\nabla^n \mathcal{D}_{t,m}^\ell \mathbf{b}_m\|_{L^\infty(\mathbb{R} \times \mathbb{R}^2)} + \|\nabla^{n-1} \mathcal{D}_{t,m}^\ell \nabla \mathbf{b}_m\|_{L^\infty(\mathbb{R} \times \mathbb{R}^2)} \quad \forall n, \ell \in \mathbb{N}_0 \text{ such that } 1 \leq n + \ell \leq N_*. \quad (2.73)$$

Here and throughout the rest of the paper we use the notation  $\mathcal{D}_{t,m}$  for the material derivative along the vector field  $\mathbf{b}_m$ , which is the scalar differential operator defined by

$$\mathcal{D}_{t,m} = \partial_t + \mathbf{b}_m \cdot \nabla. \quad (2.73)$$

Note that as opposed to (2.72), in which the index of the space derivatives is allowed to be arbitrarily large ( $n \in \mathbb{N}$ ), in (2.76) we only are concerned with a total derivative index  $n + \ell \leq N_*$  which is finite (in particular, bounded independently of  $m$ ). As such, the implicit constants in these estimates are allowed to depend on  $n$  and  $\ell$ , because this just means that they depend on  $N_*$ , and so they depend on our choice of  $\beta$  (via (2.6)). The advantage of this relaxation is that we do not need to keep track of factorial terms (e.g.  $n!, \ell!, (n+\ell)!$ ), or on powers of constants (e.g.  $C^n, C^\ell$ ). In particular, as opposed to the previous section, where we had to carefully apply the auxiliary lemmas from Appendix B.2, here we can apply standard consequences of the Leibniz and chain rules, such as

$$\|\nabla^n (f g)\|_{L^\infty} \leq C \|\nabla^n f\|_{L^\infty} \|g\|_{L^\infty} + C \|f\|_{L^\infty} \|\nabla^n g\|_{L^\infty}, \quad (2.74)$$

$$\|\nabla^n (f \circ g)\|_{L^\infty} \leq C \|\nabla f\|_{L^\infty} \|\nabla g\|_{C^{n-1}} + C \|\nabla f\|_{C^{n-1}} \|\nabla g\|_{L^\infty}^n, \quad (2.75)$$

for  $1 \leq n \leq N_*$ , where  $C$  only depends on  $n$  (hence on  $N_*$ , hence on  $\beta$ ).

The main result of this section is:

**Proposition 2.5.** *Assume that  $n, \ell \in \mathbb{N}_0$  are such that  $1 \leq n + \ell \leq N_*$ . Then, we have that*

$$\|\nabla^n \mathcal{D}_{t,m}^\ell \mathbf{b}_m\|_{L^\infty(\mathbb{R} \times \mathbb{R}^2)} \leq C \varepsilon_m^{\beta-1} (\varepsilon_m^{\beta-2})^\ell (\varepsilon_m^{-1})^n \quad (2.76)$$

$$\|\nabla^{n-1} \mathcal{D}_{t,m}^\ell \nabla \mathbf{b}_m\|_{L^\infty(\mathbb{R} \times \mathbb{R}^2)} \leq C \varepsilon_m^{\beta-2} (\varepsilon_m^{\beta-2})^\ell (\varepsilon_m^{-1})^{n-1} \quad (2.77)$$

where  $C$  only depends on  $\beta$ .

*Proof of Proposition 2.5.* We first establish (2.76). The bound (2.76) for  $\ell = 0$  was already established in (2.72). We next consider the case  $\ell = 1$ , which is the first interesting case; the proof of this case contains all the main ideas, but without the messy details about commutators.

We prove (2.76) for  $\ell = 1$  by induction on  $m$ . When  $m = 0$ , then  $\mathbf{b}_0 = 0$  by (2.35), so there is nothing to prove. Inductively, let  $m \geq 1$  and assume that (2.76) for  $\ell = 1$  holds with  $m$  replaced by  $m' \leq m-1$ . Note that the  $m$  dependence appears both through the function whose derivatives we study, namely  $\mathbf{b}_m$ , but also through the differential operator  $\mathcal{D}_{t,m}$  defined in (2.73). This nonlinear dependence on  $m$  makes it convenient to introduce the notation  $\mathbf{v}_m$  to denote the “fast part” of the vector field  $\mathbf{b}_m$ , namely

$$\mathbf{v}_m = \mathbf{b}_m - \mathbf{b}_{m-1} = \nabla^\perp \sum_{k \in \mathbb{Z}} \hat{\zeta}_{m,l_k} \zeta_{m,k} \psi_{m,k} \circ X_{m-1,l_k}^{-1}. \quad (2.78)$$

With this notation (2.73) becomes

$$\mathcal{D}_{t,m} = \mathcal{D}_{t,m-1} + \mathbf{v}_m \cdot \nabla,$$

and the bound (2.39) may be recast as

$$\|\nabla^n \mathbf{v}_m\|_{L^\infty(\mathbb{R} \times \mathbb{R}^2)} \leq C \varepsilon_m^{\beta-1} (n-1)! (C \varepsilon_m^{-1})^n \quad (2.79)$$

for all  $n \in \mathbb{N}_0$ . The difference with (2.72) is that (2.79) includes the case  $n = 0$ .

Next, we note that

$$\mathcal{D}_{t,m} \mathbf{b}_m - \mathcal{D}_{t,m-1} \mathbf{b}_{m-1} = \mathcal{D}_{t,m-1} \mathbf{v}_m + (\mathbf{v}_m \cdot \nabla) \mathbf{b}_{m-1} + (\mathbf{v}_m \cdot \nabla) \mathbf{v}_m. \quad (2.80)$$

The reason for the above decomposition lies in the fact that the term  $\mathcal{D}_{t,m-1} \mathbf{v}_m$  contains an important cancellation, namely  $\mathcal{D}_{t,m-1} X_{m-1,l_k}^{-1} = 0$ , and as such

$$\begin{aligned} \mathcal{D}_{t,m-1} \mathbf{v}_m &= [\mathcal{D}_{t,m-1}, \nabla^\perp] \sum_{k \in \mathbb{Z}} \hat{\zeta}_{m,l_k} \zeta_{m,k} \psi_{m,k} \circ X_{m-1,l_k}^{-1} + \nabla^\perp \mathcal{D}_{t,m-1} \sum_{k \in \mathbb{Z}} \hat{\zeta}_{m,l_k} \zeta_{m,k} \psi_{m,k} \circ X_{m-1,k}^{-1} \\ &= -\nabla^\perp \mathbf{b}_{m-1} \cdot \nabla \sum_{k \in \mathbb{Z}} \hat{\zeta}_{m,l_k} \zeta_{m,k} \psi_{m,k} \circ X_{m-1,l_k}^{-1} + \nabla^\perp \sum_{k \in \mathbb{Z}} \partial_t(\hat{\zeta}_{m,l_k} \zeta_{m,k}) \psi_{m,k} \circ X_{m-1,l_k}^{-1} \\ &= -\nabla^\perp \mathbf{b}_{m-1}^\perp \cdot \mathbf{v}_m + \nabla^\perp \sum_{k \in \mathbb{Z}} \partial_t(\hat{\zeta}_{m,l_k} \zeta_{m,k}) \psi_{m,k} \circ X_{m-1,l_k}^{-1}. \end{aligned} \quad (2.81)$$

Identity (2.81) makes formal the intuition that the “cost” of  $\mathcal{D}_{t,m-1}$  acting on  $\mathbf{v}_m$  is equal to the maximum between  $\|\nabla \mathbf{b}_{m-1}\|_{L_{t,x}^\infty}$  and  $\|\partial_t(\hat{\zeta}_{m,l_k} \zeta_{m,k})\|_{L_t^\infty}$ . Indeed, from (2.16), (2.25), (2.31), (2.72) (with  $n = 1$ ), and (2.79) (with  $n = 0$ ) we deduce from (2.81) that

$$\|\mathcal{D}_{t,m-1} \mathbf{v}_m\|_{L^\infty(\mathbb{R} \times \mathbb{R}^2)} \leq C(\varepsilon_{m-1}^{\beta-2} + \tau_m^{-1}) \varepsilon_m^{\beta-1} \leq C \tau_m^{-1} \varepsilon_m^{\beta-1}.$$

In fact, by appealing to (2.21), (2.40), (2.72), (2.74), (2.75), and (2.79), we deduce from (2.81) that for  $0 \leq n \leq N_* - 1$ ,

$$\|\nabla^n \mathcal{D}_{t,m-1} \mathbf{v}_m\|_{L^\infty(\mathbb{R} \times \mathbb{R}^2)} \leq C \tau_m^{-1} \varepsilon_m^{\beta-1} (\varepsilon_{m-1}^{-n} + \varepsilon_m^{-n} + \varepsilon_m^{-1} \varepsilon_{m-1}^{-n+1} + \varepsilon_m^{-n}) \leq C \tau_m^{-1} \varepsilon_m^{\beta-1} \varepsilon_m^{-n}. \quad (2.82)$$

This handles the estimates for the first term on the right side of (2.80). For the second term on the right side of (2.80), by (2.72), (2.79), and the Leibniz rule, we deduce that

$$\|\nabla^n (\mathbf{v}_m \cdot \nabla \mathbf{b}_{m-1})\|_{L^\infty(\mathbb{R} \times \mathbb{R}^2)} \leq C \varepsilon_m^{\beta-1} \varepsilon_{m-1}^{\beta-2} (\varepsilon_m^{-n} + \varepsilon_{m-1}^{-n}) \leq C \varepsilon_m^{\beta-1} \tau_m^{-1} \varepsilon_m^{-n}. \quad (2.83)$$

Also, by (2.79) and the Leibniz rule we obtain

$$\|\nabla^n (\mathbf{v}_m \cdot \nabla \mathbf{v}_m)\|_{L^\infty(\mathbb{R} \times \mathbb{R}^2)} \leq C \varepsilon_m^{\beta-1} \varepsilon_m^{\beta-2} \varepsilon_m^{-n}. \quad (2.84)$$

Comparing the bounds in (2.82)–(2.84), and noting that (2.16) and (2.5) give

$$\tau_m^{-1} \leq C \varepsilon_{m-1}^{\beta-2-4\delta} \leq C \varepsilon_{m-1}^{q(\beta-2)} \leq C \varepsilon_m^{\beta-2} \quad (2.85)$$

we deduce from (2.80) that for all  $0 \leq n \leq N_* - 1$ , we have that

$$\|\nabla^n (\mathcal{D}_{t,m} \mathbf{b}_m - \mathcal{D}_{t,m-1} \mathbf{b}_{m-1})\|_{L^\infty(\mathbb{R} \times \mathbb{R}^2)} \leq C \varepsilon_m^{\beta-1} \varepsilon_m^{\beta-2} \varepsilon_m^{-n}.$$

Lastly, using that  $(\beta - 2) + (\beta - 1) - n \leq 2\beta - 3 \leq 8/3 - 3 = -1/3 < 0$  and  $\varepsilon_m \leq \varepsilon_{m-1}$ , we may use that by induction the estimate (2.76) holds at level  $m - 1$ , to deduce

$$\|\nabla^n \mathcal{D}_{t,m} \mathbf{b}_m\|_{L^\infty(\mathbb{R} \times \mathbb{R}^2)} \leq C \varepsilon_m^{\beta-2} \varepsilon_m^{\beta-1} \varepsilon_m^{-n} + C \varepsilon_{m-1}^{\beta-2} \varepsilon_{m-1}^{\beta-1} \varepsilon_{m-1}^{-n} \leq C \varepsilon_m^{2\beta-3-n}, \quad (2.86)$$

where the constant  $C$  is independent of  $m$ , but may depend on  $n \leq N_* - 1$ . This establishes (2.76) at level  $m$ , when  $\ell = 1$ . By induction on  $m$ , we have proven the bound (2.76) for  $\ell = 1$  and  $m \geq 0$ .

The proof of (2.76) for  $\ell \geq 2$  proceeds in a similar manner, but it requires a number of commutator estimates, because the operators  $\{\mathcal{D}_{t,m}, \mathbf{v}_m \cdot \nabla, \nabla, \partial_t\}$ , do not commute. As before, when  $m = 0$  there is nothing to prove because  $\mathbf{b}_0 = 0$ . Inductively, let  $m \geq 1$  and assume that (2.76) for  $\ell = 2$  holds with  $m$  replaced by  $m' \leq m - 1$ . By differentiating (2.80) with respect to  $\mathcal{D}_{t,m} = \mathcal{D}_{t,m-1} + \mathbf{v}_m \cdot \nabla$ , we obtain

$$\begin{aligned} \mathcal{D}_{t,m}^2 \mathbf{b}_m - \mathcal{D}_{t,m-1}^2 \mathbf{b}_{m-1} - \mathcal{D}_{t,m-1}^2 \mathbf{v}_m \\ = (\mathbf{v}_m \cdot \nabla) \mathcal{D}_{t,m-1} \mathbf{b}_{m-1} + (\mathbf{v}_m \cdot \nabla) \mathcal{D}_{t,m-1} \mathbf{v}_m + (\mathcal{D}_{t,m-1} \mathbf{v}_m \cdot \nabla) \mathbf{b}_m \\ + (\mathbf{v}_m \cdot \nabla) \mathcal{D}_{t,m} \mathbf{b}_m - ((\mathbf{v}_m \cdot \nabla) \mathbf{b}_{m-1} \cdot \nabla) \mathbf{b}_m. \end{aligned} \quad (2.87)$$

All terms on the right side of (2.87) contain at most one material derivative, and therefore are already bounded in light of (2.79), (2.82), and (2.76) with  $\ell \in \{0, 1\}$ . By also appealing to (2.74) and (2.85), we obtain for all  $n \leq N_* - 2$ ,

$$\begin{aligned} \|\nabla^n (\text{RHS of (2.87)})\|_{L^\infty(\mathbb{R} \times \mathbb{R}^2)} &\leq C \varepsilon_m^{\beta-1} \varepsilon_{m-1}^{2(\beta-2)} (\varepsilon_m^{-n} + \varepsilon_{m-1}^{-n}) + C \varepsilon_m^{\beta-1} \tau_m^{-1} \varepsilon_m^{\beta-2} \varepsilon_m^{-n} \\ &\quad + C \varepsilon_m^{\beta-1} \varepsilon_m^{2(\beta-2)} \varepsilon_m^{-n} + C \varepsilon_m^{\beta-1} \varepsilon_{m-1}^{\beta-2} \varepsilon_m^{\beta-2} (\varepsilon_m^{-n} + \varepsilon_{m-1}^{-n}) \\ &\leq C \varepsilon_m^{\beta-1} (\varepsilon_m^{\beta-2})^2 \varepsilon_m^{-n}. \end{aligned} \quad (2.88)$$

In order to estimate the contribution from  $\mathcal{D}_{t,m-1}^2 \mathbf{v}_m$ , we apply  $\mathcal{D}_{t,m-1}$  to (2.81), and deduce that

$$\begin{aligned} \mathcal{D}_{t,m-1}^2 \mathbf{v}_m &= -\nabla^\perp \mathbf{b}_{m-1}^\perp \cdot \mathcal{D}_{t,m-1} \mathbf{v}_m - \nabla^\perp \mathcal{D}_{t,m-1} \mathbf{b}_{m-1}^\perp \cdot \mathbf{v}_m + \nabla^\perp \mathbf{b}_{m-1} \cdot \nabla \mathbf{b}_{m-1}^\perp \cdot \mathbf{v}_m \\ &\quad - \nabla^\perp \mathbf{b}_{m-1}^\perp \cdot \nabla^\perp \sum_{k \in \mathbb{Z}} \partial_t (\hat{\zeta}_{m,l_k} \zeta_{m,k}) \psi_{m,k} \circ X_{m-1,l_k}^{-1} + \nabla^\perp \sum_{k \in \mathbb{Z}} \partial_t^2 (\hat{\zeta}_{m,l_k} \zeta_{m,k}) \psi_{m,k} \circ X_{m-1,l_k}^{-1}. \end{aligned}$$

Thus, by appealing to (2.21), (2.25), (2.31), (2.40), (2.72), (2.74), (2.76) with  $\ell = 1$ , (2.79), (2.82), and (2.85), analogously to (2.88) we may deduce that for  $n \leq N_* - 2$ ,

$$\|\nabla^n \mathcal{D}_{t,m-1}^2 \mathbf{v}_m\|_{L^\infty(\mathbb{R} \times \mathbb{R}^2)} \leq C \varepsilon_m^{\beta-1} (\varepsilon_m^{\beta-2})^2 \varepsilon_m^{-n}. \quad (2.89)$$

Combining (2.87), (2.88), (2.89), and the inductive assumption that (2.76) holds for  $\ell = 2$  at level  $m - 1$ , we deduce that for  $n \leq N_* - 2$ ,

$$\begin{aligned} \|\nabla^n \mathcal{D}_{t,m}^2 \mathbf{b}_m\|_{L^\infty(\mathbb{R} \times \mathbb{R}^2)} &\leq \|\nabla^n \mathcal{D}_{t,m-1}^2 \mathbf{b}_{m-1}\|_{L^\infty(\mathbb{R} \times \mathbb{R}^2)} + C \varepsilon_m^{\beta-1} (\varepsilon_m^{\beta-2})^2 \varepsilon_m^{-n} \\ &\leq C \varepsilon_{m-1}^{\beta-1} (\varepsilon_{m-2}^{\beta-2})^2 \varepsilon_{m-1}^{-n} + C \varepsilon_m^{\beta-1} (\varepsilon_m^{\beta-2})^2 \varepsilon_m^{-n} \\ &\leq C \varepsilon_m^{\beta-1} (\varepsilon_m^{\beta-2})^2 \varepsilon_m^{-n}. \end{aligned}$$

By induction on  $m$ , this concludes the proof of (2.76) for  $\ell = 2$ .

Estimate (2.76) in the case  $3 \leq \ell \leq N_*$  and  $0 \leq n \leq N_* - \ell$  may be proven in the same way as for  $\ell \in \{1, 2\}$ , save for the bookkeeping, which becomes tedious. Inductively on  $\ell$ , in analogy to (2.80) and (2.87) we may show that the expression  $\mathcal{D}_{t,m}^\ell \mathbf{b}_m - \mathcal{D}_{t,m-1}^\ell \mathbf{b}_{m-1} - \mathcal{D}_{t,m-1}^\ell \mathbf{v}_m$  is given by a

sum of terms which contain at most  $\ell - 1$  material derivatives, and are hence bounded by induction. The precise accounting of all terms requires estimates for high-order commutators with material derivatives (e.g.  $[\nabla^n, \mathcal{D}_{t,m-1}^\ell]$ ), and for powers of sums of non-commuting operators (e.g.  $(\mathcal{D}_{t,m-1} + \mathbf{v}_m \cdot \nabla)^\ell$ ). Such estimates are given in [BMNV23, Appendices A.6 and A.7]. Using [BMNV23, Appendices A.6 and A.7], we may show that every additional material derivative  $\mathcal{D}_{t,m}$  landing on  $\mathbf{b}_m$  “costs” a factor of at most  $\varepsilon_m^{\beta-2}$ , while every additional space derivatives “costs” a factor of at most  $\varepsilon_m^{-1}$ . We omit these details.

We now turn to the proof of (2.77). This bound follows from (2.76) if we are able to estimate the commutator  $[\mathcal{D}_{t,m}, \nabla] \mathbf{b}_m$ . When  $\ell = 1$ , this commutator equals  $[\mathcal{D}_{t,m}, \nabla] = -\nabla \mathbf{b}_m \cdot \nabla$ , and hence

$$\nabla^{n-1} \mathcal{D}_{t,m} \nabla \mathbf{b}_m = \nabla^n \mathcal{D}_{t,m} \mathbf{b}_m - \nabla^{n-1} (\nabla \mathbf{b}_m \cdot \nabla \mathbf{b}_m).$$

Upon appealing to (2.72), (2.74), and (2.76) we obtain

$$\|\nabla^{n-1} \mathcal{D}_{t,m} \nabla \mathbf{b}_m\|_{L^\infty(\mathbb{R} \times \mathbb{R}^2)} \leq C \varepsilon_m^{\beta-1} (\varepsilon_m^{\beta-2}) (\varepsilon_m^{-1})^n + C (\varepsilon_m^{\beta-2})^2 (\varepsilon_m^{-1})^{n-1} = C (\varepsilon_m^{\beta-2})^2 (\varepsilon_m^{-1})^{n-1}.$$

This establishes (2.77) when  $\ell = 1$ , for  $n \leq N_* - 1$ .

In order to prove (2.77) for  $\ell \geq 2$ , we assume by induction that (2.77) holds for  $\ell' \leq \ell - 1$ . At this stage, we recall from [BMNV23, Lemma A.12] that the commutator between high powers of the material derivative operator and a space-gradient is given by

$$[\mathcal{D}_{t,m}^\ell, \nabla] = \sum_{\ell'=1}^{\ell} \binom{\ell}{\ell'} (\text{ad} \mathcal{D}_{t,m})^{\ell'} (\nabla) \mathcal{D}_{t,m}^{\ell-\ell'}, \quad (2.90)$$

where  $(\text{ad} \mathcal{D}_{t,m})^0(\nabla) = \nabla$ , and recursively we define  $(\text{ad} \mathcal{D}_{t,m})^{\ell'}(\nabla) = [\mathcal{D}_{t,m}, (\text{ad} \mathcal{D}_{t,m})^{\ell'-1}(\nabla)]$ . In turn, using [BMNV23, Lemma A.13], we have that

$$(\text{ad} \mathcal{D}_{t,m})^\ell(\nabla) = \sum_{j=1}^{\ell} \sum_{\alpha \in \mathbb{N}_0^j : |\alpha|=\ell-j} c_{\ell,j,\alpha} \prod_{i=1}^j (\mathcal{D}_{t,m}^{\alpha_i} \nabla \mathbf{b}_m) \cdot \nabla, \quad (2.91)$$

where the product  $\prod_{i=1}^j$  is the product of matrices, and the coefficients  $c_{\ell,j,\alpha}$  only depend on  $\ell, j, \alpha$ . From (2.90) and (2.91) we thus obtain

$$\begin{aligned} \|\nabla^{n-1} \mathcal{D}_{t,m}^\ell \nabla \mathbf{b}_m\|_{L_{t,x}^\infty} &\leq \|\nabla^n \mathcal{D}_{t,m}^\ell \mathbf{b}_m\|_{L_{t,x}^\infty} + \|\nabla^{n-1} [\mathcal{D}_{t,m}^\ell, \nabla] \mathbf{b}_m\|_{L_{t,x}^\infty} \\ &\leq \|\nabla^n \mathcal{D}_{t,m}^\ell \mathbf{b}_m\|_{L_{t,x}^\infty} + C \sum_{n'=0}^{n-1} \sum_{\ell'=1}^{\ell} \|\nabla^{n'} (\text{ad} \mathcal{D}_{t,m})^{\ell'}(\nabla) \nabla^{n-1-n'} \mathcal{D}_{t,m}^{\ell-\ell'} \mathbf{b}_m\|_{L_{t,x}^\infty}. \end{aligned}$$

Since  $\ell - \ell' \leq \ell - 1$ , and upon noting that the  $\alpha$  in (2.91) satisfies  $|\alpha| = \ell - j \leq \ell - 1$ , we deduce that from the inductive assumption (2.77) and the bound (2.76) that

$$\begin{aligned} \|\nabla^{n-1} \mathcal{D}_{t,m}^\ell \nabla \mathbf{b}_m\|_{L_{t,x}^\infty} &\leq C \varepsilon_m^{\beta-1} (\varepsilon_m^{\beta-2})^\ell (\varepsilon_m^{-1})^n \\ &\quad + C \sum_{n'=0}^{n-1} \sum_{\ell'=1}^{\ell} \sum_{j=1}^{\ell'} \sum_{\alpha \in \mathbb{N}_0^j : |\alpha|=\ell'-j} \|\nabla^{n'} \prod_{i=1}^j (\mathcal{D}_{t,m}^{\alpha_i} \nabla \mathbf{b}_m)\|_{L_{t,x}^\infty} \|\nabla^{n-n'} \mathcal{D}_{t,m}^{\ell-\ell'} \mathbf{b}_m\|_{L_{t,x}^\infty} \\ &\leq C \varepsilon_m^{\beta-1} (\varepsilon_m^{\beta-2})^\ell (\varepsilon_m^{-1})^n \\ &\quad + C \sum_{n'=0}^{n-1} \sum_{\ell'=1}^{\ell} \sum_{j=1}^{\ell'} \sum_{\alpha \in \mathbb{N}_0^j : |\alpha|=\ell'-j} (\varepsilon_m^{\beta-2})^{|\alpha|+j} (\varepsilon_m^{-1})^{n'} \varepsilon_m^{\beta-1} (\varepsilon_m^{\beta-2})^{\ell-\ell'} (\varepsilon_m^{-1})^{n-n'} \\ &\leq C \varepsilon_m^{\beta-1} (\varepsilon_m^{\beta-2})^\ell (\varepsilon_m^{-1})^n. \end{aligned}$$

A close inspection of the above chain of inequalities reveals that the total number of space, plus the total number of material derivatives, never exceeds  $N_*$  as desired. This establishes the inductive step for (2.77), concluding the proof.  $\square$

An immediate consequence of Proposition 2.5 is an estimate for  $\nabla^n \partial_t^\ell$  applied to  $\nabla X_{m,l}^{-1} \circ X_{m,l}$ .

**Corollary 2.6.** *Assume that  $n, \ell \in \mathbb{N}_0$  are such that  $n + \ell \leq N_*$ . Then, for all  $t, s \in \mathbb{R}$  with  $|t| \leq 2^{-25} a_m^{-1}$ , we have that*

$$\|\nabla^n \partial_t^\ell (\nabla X_m^{-1}(t+s, X_m(t+s, \cdot, s), s))\|_{L^\infty(\mathbb{R}^2)} \leq C \varepsilon_m^{-n} (\varepsilon_m^{\beta-2})^\ell, \quad (2.92)$$

where the constant  $C \geq 1$  depends only on  $\beta$ , through  $N_*$ .

*Proof of Corollary 2.6.* When  $\ell = 0$  and  $n = 0$ , the bound (2.92) follows from (2.40), while for  $\ell = 0$  and  $1 \leq n \leq N_*$  we additionally appeal to (2.67), (2.70), and (2.75) to deduce

$$\begin{aligned} \|\nabla^n (\nabla X_m^{-1} \circ X_m)\|_{L^\infty(\mathbb{R}^2)} &\leq C \|\nabla^2 X_m^{-1}\|_{L^\infty} \|\nabla X_m\|_{C^{n-1}} + C \|\nabla^2 X_m^{-1}\|_{C^{n-1}} \|\nabla X_m\|_{L^\infty}^n \\ &\leq C \varepsilon_m^{-1} (\varepsilon_m^{-1})^{n-1} + C (\varepsilon_m^{-1})^n = C \varepsilon_m^{-n}. \end{aligned}$$

As such, it only remains to prove (2.92) for  $\ell \geq 1$  and  $n \leq N_* - \ell$ .

Using that  $\mathcal{D}_{t,m} X_m^{-1} = 0$ , for  $1 \leq \ell \leq N_*$  we have

$$\partial_t^\ell (\nabla X_m^{-1} \circ X_m) = (\mathcal{D}_{t,m}^\ell \nabla X_m^{-1}) \circ X_m = ([\mathcal{D}_{t,m}^\ell, \nabla] X_m^{-1}) \circ X_m. \quad (2.93)$$

The formula for the commutator present in (2.93) was recorded earlier in (2.90). By again using that  $\mathcal{D}_{t,m} X_m^{-1} = 0$ , we deduce from (2.90) that

$$[\mathcal{D}_{t,m}^\ell, \nabla] X_m^{-1} = \sum_{\ell'=1}^{\ell} \binom{\ell}{\ell'} (\text{ad } \mathcal{D}_{t,m})^{\ell'} (\nabla) \mathcal{D}_{t,m}^{\ell-\ell'} X_m^{-1} = (\text{ad } \mathcal{D}_{t,m})^\ell (\nabla) X_m^{-1}, \quad (2.94)$$

where the first order differential operator  $(\text{ad } \mathcal{D}_{t,m})^\ell (\nabla)$  is given by (2.91). With the available bound (2.77), and the identity (2.91), we bound the left side of (2.94) as

$$\begin{aligned} &\|\nabla^n [\mathcal{D}_{t,m}^\ell, \nabla] X_m^{-1}\|_{L_{t,x}^\infty} \\ &= \|\nabla^n (\text{ad } \mathcal{D}_{t,m})^\ell (\nabla) X_m^{-1}\|_{L_{t,x}^\infty} \\ &\leq C \sum_{n'=0}^n \sum_{j=1}^{\ell} \sum_{\alpha \in \mathbb{N}_0^j: |\alpha|=\ell-j} \sum_{\beta \in \mathbb{N}_0^j: |\beta|=n'} \prod_{i=1}^j \|\nabla^{\beta_i} \mathcal{D}_{t,m}^{\alpha_i} \nabla \mathbf{b}_m\|_{L_{t,x}^\infty} \|\nabla^{n-n'+1} X_m^{-1}\|_{L_{t,x}^\infty} \\ &\leq C \sum_{n'=0}^n \sum_{j=1}^{\ell} \sum_{\alpha \in \mathbb{N}_0^j: |\alpha|=\ell-j} \sum_{\beta \in \mathbb{N}_0^j: |\beta|=n'} \prod_{i=1}^j (\varepsilon_m^{\beta-2})^{\alpha_i+1} (\varepsilon_m^{-1})^{\beta_i} (\varepsilon_m^{-1})^{n-n'} \\ &\leq C (\varepsilon_m^{\beta-2})^\ell (\varepsilon_m^{-1})^n. \end{aligned} \quad (2.95)$$

In order to conclude the proof of (2.92), we combine the above estimate with the identity (2.93), and (2.75), to obtain

$$\begin{aligned} &\|\nabla^n \partial_t^\ell (\nabla X_m^{-1} \circ X_m)\|_{L_{t,x}^\infty} \\ &\leq C \|\nabla [\mathcal{D}_{t,m}^\ell, \nabla] X_m^{-1}\|_{L_{t,x}^\infty} \|\nabla X_m\|_{L_t^\infty C_x^{n-1}} + C \|\nabla [\mathcal{D}_{t,m}^\ell, \nabla] X_m^{-1}\|_{L_t^\infty C_x^{n-1}} \|\nabla X_m\|_{L_{t,x}^\infty}^n \\ &\leq C (\varepsilon_m^{\beta-2})^\ell (\varepsilon_m^{-1})^n. \end{aligned}$$

In the above estimate, the time-support was not written out explicitly, but it was implicitly assumed to be such that  $|t| \leq 2^{-25}a_m^{-1}$ , so that we could appeal to the bounds (2.40) and (2.70). This concludes the proof of the corollary.  $\square$

A second consequence of Proposition 2.5 is the following estimate on  $\nabla^n \mathcal{D}_{t,m}^\ell(\nabla X_{m,l} \circ X_{m,l}^{-1})$ . This is necessary to estimate the spatial and material derivatives of the matrix  $\mathbf{s}_m$  in Section 4.3.

**Corollary 2.7.** *Assume that  $n, \ell \in \mathbb{N}_0$  are such that  $n + \ell \leq N_*$ . Then, for all  $t, s \in \mathbb{R}$  with  $|t| \leq 2^{-25}a_m^{-1}$ , we have that*

$$\|\nabla^n \mathcal{D}_{t,m}^\ell(\nabla X_m(t + s, X_m^{-1}(t + s, \cdot, s), s))\|_{L^\infty(\mathbb{R}^2)} \leq C\varepsilon_m^{-n}(\varepsilon_m^{\beta-2})^\ell, \quad (2.96)$$

where the constant  $C \geq 1$  depends only on  $\beta$ , through  $N_*$ .

*Proof of Corollary 2.7.* When  $\ell = 0$ , the desired bounds were already obtained in (2.68). For  $\ell \geq 1$ , the proof of (2.96) starts with the observation that  $\nabla X_m \circ X_m^{-1} = (\nabla X_m^{-1})^{-1}$  as  $2 \times 2$  matrices. Since the flows  $\mathbf{b}_m$  that define  $X_m$  are incompressible, we have that  $\det(\nabla X_m^{-1}) = 1$ , and so  $\nabla X_m \circ X_m^{-1}$  equals the transpose of the cofactor matrix associated to  $\nabla X_m^{-1}$ . In turn, since we are in two space dimensions, this cofactor matrix equals to  $\nabla^\perp(X_m^{-1})^\perp$ . This leads to the identity

$$\nabla X_m \circ X_m^{-1} = (\nabla X_m^{-1})^{-1} = (\nabla^\perp(X_m^{-1})^\perp)^T.$$

The purpose of the above identity is to show that if we have estimates for all the entries of the matrix  $\nabla^n \mathcal{D}_{t,m}^\ell(\nabla X_m^{-1})$ , then we automatically obtain estimates for the matrix  $\nabla^n \mathcal{D}_{t,m}^\ell(\nabla X_m \circ X_m^{-1})$ .

To conclude, we note that since  $\mathcal{D}_{t,m}^\ell X_m^{-1} = 0$ , we have that  $\nabla^n \mathcal{D}_{t,m}^\ell \nabla X_m^{-1} = \nabla^n [\mathcal{D}_{t,m}^\ell, \nabla] X_m^{-1}$ , and precisely this term was previously estimated in (2.95). In turn, this estimate recovers (2.96), as desired.  $\square$

We conclude this section by noting that by construction, the vector field  $\mathbf{b}$  defined in (2.38) is “nearly a solution” of the incompressible Euler equations, as quantified by the following result.

**Proposition 2.8.** *The vector field  $\mathbf{b}$  constructed in (2.38) solves*

$$\partial_t \mathbf{b} + \operatorname{div}(\mathbf{b} \otimes \mathbf{b}) + \nabla p = \operatorname{div} \mathring{\mathbf{R}}, \quad \operatorname{div} \mathbf{b} = 0, \quad (2.97)$$

for a suitable pressure scalar  $p \in C(\mathbb{R}; C^{0,\beta'}(\mathbb{R}^2))$  and a traceless stress tensor  $\mathring{\mathbf{R}} \in C(\mathbb{R}; C^{0,\beta'}(\mathbb{R}^2))$ , for any  $\beta' \in (0, 1)$  satisfying

$$\beta' < 2(\beta - 1) + \frac{2\delta}{q} = 2(\beta - 1) + \frac{(\beta - 1)(4 - 3\beta)^2}{2\beta(5\beta - 4)}. \quad (2.98)$$

Moreover, there exists a constant  $C \geq 1$  which only depends on  $\beta$  and  $\beta'$  as in (2.98), such that

$$\|p\|_{L^\infty(\mathbb{R}; C^{0,\beta'}(\mathbb{R}^2))} + \|\mathring{\mathbf{R}}\|_{L^\infty(\mathbb{R}; C^{0,\beta'}(\mathbb{R}^2))} \leq C\Lambda^{-\frac{q}{q-1}(2(\beta-1)-\beta')} + C\Lambda^{-\frac{2q}{q-1}(2(\beta-1)+\frac{2\delta}{q}-\beta')}. \quad (2.99)$$

In particular, for  $\beta, \beta'$  fixed, the right side of (2.99) can be made arbitrarily small by letting  $\Lambda$  be sufficiently large.

We note that the parameter  $\beta'$  in (2.98) is allowed to be strictly larger than  $2(\beta - 1)$ . As such the regularity of the pressure in (2.97) is strictly better than the regularity of the pressure for a generic  $C_t^0 C_x^{0,\beta-1}$  weak solution of the Euler equations, which is  $C_t^0 C_x^{0,2(\beta-1)}$ . Proposition 2.8 follows from a fairly straightforward computation by telescoping (2.80) and (2.81) and the fact that the term  $\mathbf{v}_m \cdot \nabla \mathbf{v}_m$  vanishes to leading order due to the shear flows used in the construction. We do not give the details here, since it is not needed in our analysis, but the interested reader can find the proof commented out in the latex source file (downloadable on the arxiv) below this sentence.

### 3. Correctors and renormalized diffusivities

In this section, we introduce the sequence of correctors and renormalized diffusivities for each scale  $\varepsilon_m$ .

#### 3.1 The correctors: definitions and estimates

We will introduce a corrector  $\chi_m^\kappa$  which mediates between scales  $\varepsilon_m$  and  $\varepsilon_{m-1}$ . The job of  $\chi_m^\kappa$  is to “correct” a solution of the  $\varepsilon_{m-1}$ -scale equation

$$\partial_t \theta_{m-1} - \kappa_{m-1} \Delta \theta_{m-1} + \mathbf{b}_{m-1} \cdot \nabla \theta_{m-1} = 0, \quad (3.1)$$

by adding the wiggles with wavelengths of order  $\varepsilon_m$  we would expect to see in the solution of the  $\varepsilon_m$ -scale equation.

As we have seen in the construction of the vector field, the difference between  $\mathbf{b}_m$  and  $\mathbf{b}_{m-1}$  is the inclusion of shear flows oscillating at the length scale  $\varepsilon_m$ , in the Lagrangian coordinates corresponding to  $\mathbf{b}_{m-1}$ . These shear flows alternate between horizontal and vertical shears (with “quiet” periods in between) on the time scale  $\tau_m$  which, as we will show, is much longer than the time scale on which it takes for the shear flows to homogenize. These oscillations in space and in time will create oscillations in the solutions  $\theta_m$  which are not present in  $\theta_{m-1}$ . We need to introduce correctors which capture, at leading order, these oscillations. Roughly speaking, the correctors which capture the spatial oscillations at scale  $\varepsilon_m$  in Lagrangian coordinates will be denoted by  $\tilde{\chi}_m$ . The time oscillations due to the horizontal and vertical alternation of the shear flows will be corrected by a function denoted by  $\tilde{H}_m$ .

Since the scale  $\varepsilon_m$  on which the shears oscillate is much smaller than the active scales of the flows  $X_{m-1}$ , we should expect the correctors  $\tilde{\chi}_m$  to be obtained—at least at leading order—from the composition of the correctors for the (time-independent) simple shear flow with the appropriate Lagrangian flow  $X_{m-1}$ . As such, we first discuss the derivation of the correctors corresponding to the time-independent shear flows. We denote these by  $\chi_{m,k}$ , and they turn out to be given by an explicit, well-known and simple formula.

#### 3.2 Correctors and effective diffusivity for smoothly alternating shear flows

All of the notation from the previous section is adopted here; in particular we recall that the stream function  $\psi_{m,k}$  is defined for each  $k \in \mathbb{Z}$  and  $m \in \mathbb{N}$  in (2.19)–(2.20) and the time cutoff functions  $\zeta_{m,k}$  and  $\hat{\zeta}_{m,l}$  are defined in (2.22)–(2.24) and (2.29)–(2.31), respectively, and we recall from (2.32) that the supports of these overlap only when  $l = l_k$ , with  $l_k$  defined in (2.17). We also define the incompressible vector fields  $\mathbf{u}_{m,k}$  by

$$\mathbf{u}_{m,k}(x) := \nabla^\perp \psi_{m,k}(x) = \begin{cases} 2\pi a_m \varepsilon_m \cos\left(\frac{2\pi x_1}{\varepsilon_m}\right) \mathbf{e}_2 & \text{if } k \in 4\mathbb{Z} + 1, \\ -2\pi a_m \varepsilon_m \cos\left(\frac{2\pi x_2}{\varepsilon_m}\right) \mathbf{e}_1 & \text{if } k \in 4\mathbb{Z} + 3, \\ 0 & \text{if } k \in 2\mathbb{Z}, \end{cases} \quad (3.2)$$

and we set

$$\psi_m(t, x) := \sum_{k \in 2\mathbb{Z} + 1} \hat{\zeta}_{m,l_k}(t) \zeta_{m,k}(t) \psi_{m,k}(x). \quad (3.3)$$

and

$$\mathbf{u}_m(t, x) := \nabla^\perp \psi_m(t, x) = \sum_{k \in 2\mathbb{Z} + 1} \hat{\zeta}_{m,l_k}(t) \zeta_{m,k}(t) \mathbf{u}_{m,k}(x). \quad (3.4)$$

We let  $\nabla \mathbf{u}_m(t, x)$  denote the 2-by-2 matrix with entries  $\partial_{x_i}(e_j \cdot \mathbf{u}_m)(t, x)$ ; it is given by the formula

$$\nabla \mathbf{u}_{m,k}(x) = \begin{cases} -4\pi^2 a_m \sin\left(\frac{2\pi x_1}{\varepsilon_m}\right) \mathbf{e}_1 \otimes \mathbf{e}_2 & \text{if } k \in 4\mathbb{Z} + 1, \\ 4\pi^2 a_m \sin\left(\frac{2\pi x_2}{\varepsilon_m}\right) \mathbf{e}_2 \otimes \mathbf{e}_1 & \text{if } k \in 4\mathbb{Z} + 3, \\ 0 & \text{if } k \in 2\mathbb{Z}. \end{cases} \quad (3.5)$$

We also introduce a special time  $t_{m,k}^* := (-\frac{2}{3} + k)\tau_m$  and, for each  $\kappa > 0$  and  $\mathbf{e} \in \mathbb{R}^2$ , define  $\chi_{m,k,\mathbf{e}}^\kappa$  to be the solution of

$$\begin{cases} \partial_t \chi_{m,k,\mathbf{e}}^\kappa - \kappa \Delta \chi_{m,k,\mathbf{e}}^\kappa + \hat{\zeta}_{m,l_k} \zeta_{m,k} \mathbf{u}_{m,k} \cdot (\mathbf{e} + \nabla \chi_{m,k,\mathbf{e}}^\kappa) = 0 & \text{in } (-\infty, \infty) \times \mathbb{R}^2, \\ \chi_{m,k,\mathbf{e}}^\kappa = 0 & \text{in } (-\infty, t_{m,k}^*) \times \mathbb{R}^2. \end{cases} \quad (3.6)$$

We observe that (3.6) does have a unique solution by first imposing a zero initial condition at time  $t_{m,k}^*$  and then noticing that, thanks to the presence of the cutoff function  $\zeta_{m,k}$ , we may extend the solution to earlier times by setting it equal to zero. Note that  $\chi_{m,k}^\kappa \equiv 0$  for  $k \in 2\mathbb{Z}$  by (3.2).

We will use vector notation for these correctors by writing

$$\boldsymbol{\chi}_{m,k}^\kappa := \begin{pmatrix} \chi_{m,k,\mathbf{e}_1}^\kappa \\ \chi_{m,k,\mathbf{e}_2}^\kappa \end{pmatrix}^\top, \quad m \in \mathbb{N}.$$

Then  $\nabla \boldsymbol{\chi}_{m,k}^\kappa$  denotes the  $2 \times 2$  matrix

$$\nabla \boldsymbol{\chi}_{m,k}^\kappa = \begin{pmatrix} \partial_{x_1} \chi_{m,k,\mathbf{e}_1}^\kappa & \partial_{x_1} \chi_{m,k,\mathbf{e}_2}^\kappa \\ \partial_{x_2} \chi_{m,k,\mathbf{e}_1}^\kappa & \partial_{x_2} \chi_{m,k,\mathbf{e}_2}^\kappa \end{pmatrix}. \quad (3.7)$$

Thanks to the one-dimensional nature of the shear flows  $\mathbf{u}_{m,k}$ , we can give a simple explicit formula for  $\boldsymbol{\chi}_{m,k}^\kappa$ . Indeed, a direct computation yields

$$\boldsymbol{\chi}_{m,k}^\kappa(t, x) = \mathbf{u}_{m,k}(x) \int_{-\infty}^t \hat{\zeta}_{m,l_k}(s) \zeta_{m,k}(s) \exp\left(\frac{4\pi^2 \kappa}{\varepsilon_m^2}(s-t)\right) ds \quad (3.8)$$

and thus

$$\nabla \boldsymbol{\chi}_{m,k}^\kappa(t, x) = \nabla \mathbf{u}_{m,k}(x) \int_{-\infty}^t \hat{\zeta}_{m,l_k}(s) \zeta_{m,k}(s) \exp\left(\frac{4\pi^2 \kappa}{\varepsilon_m^2}(s-t)\right) ds. \quad (3.9)$$

Since  $\hat{\zeta}_{m,l_k} \zeta_{m,k} \leq 1$ , we have

$$\int_{-\infty}^t \hat{\zeta}_{m,l_k}(s) \zeta_{m,k}(s) \exp\left(\frac{4\pi^2 \kappa}{\varepsilon_m^2}(s-t)\right) ds \leq \int_{-\infty}^t \exp\left(\frac{4\pi^2 \kappa}{\varepsilon_m^2}(s-t)\right) ds = \frac{\varepsilon_m^2}{4\pi^2 \kappa}.$$

Therefore,

$$\|\boldsymbol{\chi}_{m,k}^\kappa\|_{L^\infty(\mathbb{R} \times \mathbb{T}^2)} + \varepsilon_m \|\nabla \boldsymbol{\chi}_{m,k}^\kappa\|_{L^\infty(\mathbb{R} \times \mathbb{T}^2)} \leq \frac{C a_m \varepsilon_m^3}{\kappa}. \quad (3.10)$$

Since  $\hat{\zeta}_{m,l_k} \leq 1$  and  $\zeta_{m,k}$  vanishes on  $[(k + \frac{2}{3})\tau_m, \infty)$ , we have, for every  $t \geq (k + \frac{3}{4})\tau_m$ ,

$$\int_{-\infty}^t \hat{\zeta}_{m,l_k}(s) \zeta_{m,k}(s) \exp\left(\frac{4\pi^2 \kappa}{\varepsilon_m^2}(s-t)\right) ds \leq \int_{-\infty}^{t - \frac{1}{12}\tau_m} \exp\left(\frac{4\pi^2 \kappa}{\varepsilon_m^2}(s-t)\right) ds = \frac{\varepsilon_m^2}{4\pi^2 \kappa} \exp\left(-\frac{\pi^2 \kappa \tau_m}{3\varepsilon_m^2}\right).$$

Therefore, after time  $(k + \frac{3}{4})\tau_m$  the corrector  $\chi_{m,k}^\kappa$  becomes exponentially small: we have

$$\begin{aligned} \sup_{t \in ((k + \frac{3}{4})\tau_m, \infty)} \left( \|\chi_{m,k}^\kappa(t, \cdot)\|_{L^\infty(\mathbb{T}^2)} + \varepsilon_m \|\nabla \chi_{m,k}^\kappa(t, \cdot)\|_{L^\infty(\mathbb{T}^2)} \right) &\leq \frac{C a_m \varepsilon_m^3}{\kappa} \exp\left(-\frac{\pi^2 \kappa \tau_m}{3 \varepsilon_m^2}\right) \\ &\leq C \varepsilon_m \exp\left(-\frac{\pi^2 \kappa \tau_m}{4 \varepsilon_m^2}\right), \end{aligned} \quad (3.11)$$

where we rather crudely used  $a_m \tau_m \leq 1$  in the last line. We will typically encounter the situation in which  $\frac{\kappa \tau_m}{\varepsilon_m^2} \gg 1$ . Indeed, it will be a negative power of  $\varepsilon_m$  in practice—see (3.46) below—and therefore the exponential factor on the right side of (3.11) is very small.

We next define

$$\chi_m^\kappa := \sum_{k \in \mathbb{Z}} \xi_{m,k} \chi_{m,k}^\kappa = \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} \chi_{m,k}^\kappa, \quad (3.12)$$

Recall that the cutoff function  $\xi_{m,k}$  is defined in (2.27) and is locally constant except for times outside the time interval  $[(k - \frac{3}{4})\tau_m, (k + \frac{3}{4})\tau_m]$ , in other words, when the function  $\chi_{m,k}^\kappa$  is very small by (3.11).

We are able to conclude by (3.11), (2.28) and superposition that the components of  $\chi_m^\kappa$  in (3.12) are “almost” solutions of

$$\begin{cases} \partial_t \chi_{m,\mathbf{e}}^\kappa - \kappa \Delta \chi_{m,\mathbf{e}}^\kappa + \mathbf{u}_m \cdot (\mathbf{e} + \nabla \chi_{m,\mathbf{e}}^\kappa) = 0 & \text{in } \mathbb{R} \times \mathbb{R}^2, \\ \chi_{m,\mathbf{e}}^\kappa(t, x) \text{ is } \mathbb{Z} \times \mathbb{Z}^2\text{-periodic,} \\ \langle \chi_{m,\mathbf{e}}^\kappa(t, \cdot) \rangle = 0, \quad \forall t \in \mathbb{R}. \end{cases} \quad (3.13)$$

We recognize (3.13) as the periodic, space-time corrector arising in parabolic homogenization. Actually, the equation in the first line of (3.13) is valid only up to an exponentially small error. Our reason for defining  $\chi_m^\kappa$  slightly differently, not in terms of (3.13) but rather as the sum (3.12), is because the exact formula (3.8) is more convenient to work with and the difference between these two is negligible. We still refer to  $\chi_m^\kappa$  as a “corrector.”

**Remark 3.1** (A special orthogonality property). An important property inherited from the shear flow structure, which will come to our rescue in the Section 5, is the following pointwise orthogonality property: for every pair of multiindices  $\alpha, \beta \in \mathbb{N}_0^d$ ,

$$\partial^\alpha \mathbf{u}_{m,k} \cdot \nabla \partial^\beta \chi_{m,k,\mathbf{e}}^\kappa = 0 \quad \text{in } \mathbb{R} \times \mathbb{T}^2. \quad (3.14)$$

Indeed, if  $k \in 4\mathbb{Z}+1$  (respectively,  $k \in 4\mathbb{Z}+3$ ), then we see from (3.8) that the function  $\partial^\beta \chi_{m,k,\mathbf{e}}^\kappa(t, \cdot)$  depends only on  $x_1$  (resp.,  $x_2$ ) and therefore its gradient is proportional to  $\mathbf{e}_1$  (resp.,  $\mathbf{e}_2$ ), while from (3.2) we see that  $\partial^\alpha \mathbf{u}_{m,k}$  is proportional to  $\mathbf{e}_2$  (resp.,  $\mathbf{e}_1$ ).

### 3.3 The renormalized diffusivities: recurrence, averaging and estimates

We introduce the following objects:

- We denote the spatially-averaged flux of the correctors by

$$\mathbf{J}_m^\kappa(t) := \left\langle (\kappa \mathbf{I}_2 + \psi_m(t, \cdot) \sigma) (\mathbf{I}_2 + \nabla \chi_m^\kappa(t, \cdot)) \right\rangle. \quad (3.15)$$

- The homogenized matrix is the average of the flux in both space and time, defined by

$$\bar{\mathbf{K}}_m^\kappa := \int_0^1 \mathbf{J}_m^\kappa(t) dt = \langle\langle (\kappa \mathbf{I}_2 + \psi_m \sigma)(\mathbf{I}_2 + \nabla \chi_m^\kappa) \rangle\rangle. \quad (3.16)$$

Since  $\chi_m$  and  $\psi_m$  are invariant under a simultaneous  $2\tau_m''\mathbb{Z}$ -translation in time and a permutation of the  $\mathbf{e}_1$  and  $\mathbf{e}_2$  axes, it follows that  $\bar{\mathbf{K}}_m^\kappa$  is a scalar matrix. We will therefore abuse notation by allowing  $\bar{\mathbf{K}}_m^\kappa$  to denote both a matrix and the positive scalar constant  $a$  such that  $\bar{\mathbf{K}}_m^\kappa = a\mathbf{I}_2$ , since it will always be clear from the context which is intended.

Using (3.8), we can find an explicit formula for  $\mathbf{J}_m^\kappa(t)$  and  $\bar{\mathbf{K}}_m^\kappa$ , which will be helpful in our computations. Observe first, using the properties of the cutoff functions, the skew-symmetry of  $\sigma$  and the fact that  $\langle \nabla \chi_{m,k}^\kappa \rangle = 0$ , that we may write (3.15) as

$$\mathbf{J}_m^\kappa(t) - \kappa \mathbf{I}_2 = \langle \psi_m \sigma \nabla \chi_m^\kappa(t, \cdot) \rangle = \left\langle \sum_{k \in 2\mathbb{Z}+1} \hat{\zeta}_{m,l_k}(t) \zeta_{m,k}(t) \psi_{m,k} \sigma \nabla \chi_{m,k}(t, \cdot) \right\rangle. \quad (3.17)$$

Using (2.20), (3.8) and  $\langle \sin^2 \rangle = \frac{1}{2}$ , we compute, for every  $k \in \mathbb{Z}$  and  $t \in [(k - \frac{1}{2})\tau_m, (k + \frac{1}{2})\tau_m]$ ,

$$\begin{aligned} & \mathbf{J}_m^\kappa(t) - \kappa \mathbf{I}_2 \\ &= 2\pi^2 a_m^2 \varepsilon_m^2 \hat{\zeta}_{m,l_k}(t) \zeta_{m,k}(t) \int_{-\infty}^t \hat{\zeta}_{m,l_k}(s) \zeta_{m,k}(s) \exp\left(\frac{4\pi^2 \kappa}{\varepsilon_m^2}(s-t)\right) ds \cdot \begin{cases} \mathbf{e}_2 \otimes \mathbf{e}_2 & \text{if } k \in 4\mathbb{Z}+1, \\ \mathbf{e}_1 \otimes \mathbf{e}_1 & \text{if } k \in 4\mathbb{Z}+3, \\ 0 & \text{if } k \in 2\mathbb{Z}. \end{cases} \end{aligned} \quad (3.18)$$

In particular, for a universal constant  $C < \infty$ ,

$$|\mathbf{J}_m^\kappa(t)| \leq \left( \kappa + \frac{C a_m^2 \varepsilon_m^4}{\kappa} \right). \quad (3.19)$$

Observe that  $\mathbf{J}_m^\kappa$  is a  $\tau_m''$ -periodic function of time. We will show next that, up to a very small error,  $\mathbf{J}_m^\kappa$  can be written as a sum of products of  $\tau_m$ -periodic functions and  $\tau_m''$ -periodic functions. Define

$$\hat{\mathbf{J}}_m^\kappa(t) := \kappa \mathbf{I}_2 + \sum_{n=0}^{N_*-1} L_{m,n}^\kappa(t) \mathbf{j}_{m,n}^\kappa(t) \quad (3.20)$$

where we define, for every  $n \in \{0, \dots, N_*\}$ ,

$$\mathbf{j}_{m,n}^\kappa(t) := \frac{2\pi^2 a_m^2 \varepsilon_m^2}{n!} \left( \frac{\varepsilon_m^2}{4\pi^2 \kappa} \right)^n \sum_{k \in 2\mathbb{Z}+1} \zeta_{m,k}(t) \partial_t^n \zeta_{m,k}(t) (\mathbf{1}_{\{k \in 4\mathbb{Z}+1\}} \mathbf{e}_2 \otimes \mathbf{e}_2 + \mathbf{1}_{\{k \in 4\mathbb{Z}+3\}} \mathbf{e}_1 \otimes \mathbf{e}_1) \quad (3.21)$$

and

$$L_{m,n}^\kappa(t) := \sum_{l \in \mathbb{Z}} \hat{\zeta}_{m,l}(t) \int_{-\infty}^t \hat{\zeta}_{m,l}(s) \left( \frac{4\pi^2 \kappa (s-t)}{\varepsilon_m^2} \right)^n \exp\left(\frac{4\pi^2 \kappa}{\varepsilon_m^2}(s-t)\right) ds. \quad (3.22)$$

Observe that  $\mathbf{j}_{m,n}^\kappa$  is indeed  $\tau_m$ -periodic, and  $L_{m,n}^\kappa$  is  $\tau_m''$ -periodic.

The functions  $\mathbf{j}_{m,n}^\kappa$  defined in (3.21) satisfy the bounds

$$\|\mathbf{j}_{m,n}^\kappa\|_{L^\infty(\mathbb{R})} \leq C a_m^2 \varepsilon_m^2 \left( \frac{\varepsilon_m^2}{\kappa \tau_m} \right)^n, \quad \forall n \in \{0, \dots, N_*\}. \quad (3.23)$$

Here the constant  $C$  depends only on  $\beta$  through  $C_{(2.22)}$ . Likewise, the functions  $L_{m,n}^\kappa$  satisfy, for every  $n, \ell \in \{0, \dots, N_*\}$ ,

$$\begin{aligned} \|\partial_t^\ell L_{m,n}^\kappa\|_{L^\infty(\mathbb{R})} &\leq \sup_{l \in \mathbb{Z}} \sup_{s \in \mathbb{R}} \|\partial_t^\ell (\hat{\zeta}_{m,l} \hat{\zeta}_{m,l}(\cdot - s))\|_{L^\infty(\mathbb{R})} \int_0^\infty \left( \frac{4\pi^2 \kappa}{\varepsilon_m^2} s \right)^n \exp\left(-\frac{4\pi^2 \kappa}{\varepsilon_m^2} s\right) ds \\ &\leq n! C^n \left( \frac{\varepsilon_m^2}{\kappa} \right) (\tau_m')^{-\ell}, \end{aligned}$$

for a constant  $C < \infty$  which depends only on  $C_{(2.31)}$  and thus only on  $\beta$ . Since  $n \leq N_*$  and  $N_*$  depends only on  $\beta$ , we deduce that, for some  $C(\beta) < \infty$ ,

$$\|\partial_t^\ell L_{m,n}^\kappa\|_{L^\infty(\mathbb{R})} \leq C \left( \frac{\varepsilon_m^2}{\kappa} \right) (\tau_m')^{-\ell}, \quad \forall \ell \in \{0, \dots, N_*\}. \quad (3.24)$$

By Taylor's formula and (2.25), for every  $s, t \in \mathbb{R}$  with  $s \leq t$  and  $N \leq N_*$ ,

$$\begin{aligned} &\left| \zeta_{m,k}(s) - \sum_{n=0}^{N_*-1} \frac{(s-t)^n}{n!} \partial_t^n \zeta_{m,k}(t) \right| \exp\left(\frac{4\pi^2 \kappa}{\varepsilon_m^2} (s-t)\right) ds \\ &\leq \frac{1}{N!} \|\partial_t^N \zeta_{m,k}\|_{L^\infty(\mathbb{R})} |s-t|^N \exp\left(\frac{4\pi^2 \kappa}{\varepsilon_m^2} (s-t)\right) ds \leq C \left( \frac{\varepsilon_m^2}{\kappa \tau_m} \right)^N \exp\left(\frac{2\pi^2 \kappa}{\varepsilon_m^2} (s-t)\right) ds. \end{aligned}$$

By the previous inequality and the triangle inequality, we obtain

$$|\mathbf{J}_m^\kappa(t) - \hat{\mathbf{J}}_m^\kappa(t)| \leq \frac{C a_m^2 \varepsilon_m^4}{\kappa} \left( \frac{\varepsilon_m^2}{\kappa \tau_m} \right)^{N_*}. \quad (3.25)$$

Since  $N_*$  is a very large constant, the estimate (3.25) says that  $\mathbf{J}_m^\kappa$  is indeed well-approximated by the function  $\hat{\mathbf{J}}_m^\kappa$ , provided that  $\varepsilon_m^2 \ll \kappa \tau_m$ .

We define a  $\tau_m''$ -periodic function  $\mathbf{K}_m^\kappa$  by averaging out the  $\tau_m$ -periodic oscillations from  $\hat{\mathbf{J}}_m^\kappa$ :

$$\mathbf{K}_m^\kappa(t) := \kappa \mathbf{I}_2 + \sum_{n=0}^{N_*-1} \langle\langle \mathbf{j}_{m,n}^\kappa \rangle\rangle L_{m,n}^\kappa(t). \quad (3.26)$$

Observe that, by (3.23) and (3.24), if  $\kappa$  satisfies

$$\varepsilon_m^2 \leq \frac{1}{2} \kappa \tau_m, \quad (3.27)$$

then we have that

$$\|\partial_t^\ell \mathbf{K}_m^\kappa\|_{L^\infty(\mathbb{R})} \leq C (\tau_m')^{-\ell} \left( \kappa + \frac{a_m^2 \varepsilon_m^4}{\kappa} \right), \quad \forall \ell \in \mathbb{N}. \quad (3.28)$$

By (3.25) and the ergodic theorem for periodic functions (Lemma C.1), we have, under the extra condition (3.27),

$$|\langle\langle \mathbf{K}_m^\kappa \rangle\rangle - \bar{\mathbf{K}}_m^\kappa| \leq \frac{C a_m^2 \varepsilon_m^4}{\kappa} \left( \frac{\varepsilon_m^2}{\kappa \tau_m} \right)^{N_*}, \quad (3.29)$$

where we recall from (3.16) that  $\bar{\mathbf{K}}_m^\kappa = \langle\langle \mathbf{J}_m^\kappa \rangle\rangle$ . Since  $N_*$  is very large, this says that  $\langle\langle \mathbf{K}_m^\kappa \rangle\rangle$  and  $\bar{\mathbf{K}}_m^\kappa$  are very close, provided that  $\varepsilon_m^2 \ll \kappa \tau_m$ .

For future reference, we introduce some higher-order time correctors for  $\mathbf{j}_{m,n}^\kappa$ : for every  $m \in \mathbb{N}$ ,  $n \in \{0, \dots, N_*\}$ , we let  $\{\mathbf{q}_{m,n,r}^\kappa\}_{r \in \mathbb{N}_0}$  be the sequence of  $\tau_m$ -periodic functions on  $\mathbb{R}$  characterized by

$$\begin{cases} \mathbf{q}_{m,n,0}^\kappa := \mathbf{j}_{m,n}^\kappa - \langle\langle \mathbf{j}_{m,n}^\kappa \rangle\rangle, \\ \langle\langle \mathbf{q}_{m,n,r}^\kappa \rangle\rangle = 0, \quad \forall r \in \mathbb{N}, \\ \partial_t \mathbf{q}_{m,n,r+1}^\kappa = -\mathbf{q}_{m,n,r}^\kappa, \quad \forall r \in \mathbb{N}. \end{cases} \quad (3.30)$$

These satisfy the bounds:

$$\|\mathbf{q}_{m,n,r}^\kappa\|_{L^\infty(\mathbb{R})} \leq C a_m^2 \varepsilon_m^2 \left( \frac{\varepsilon_m^2}{\kappa \tau_m} \right)^n \frac{(C \tau_m)^r}{r!}, \quad \forall n \in \{0, \dots, N_*\}, \quad r \in \mathbb{N}_0. \quad (3.31)$$

For reasons which will become apparent in (5.78) below, we need to compute the difference between the averaged flux  $\mathbf{J}_m^\kappa(t)$  (which is given explicitly in (3.18)) and the spatially-averaged energy

$$\mathbf{E}_m^\kappa(t) := \sum_{k,k' \in 2\mathbb{Z}+1} \xi_{m,k}(t) \xi_{m,k'}(t) \left\langle \kappa (\mathbf{I}_2 + \nabla \boldsymbol{\chi}_{m,k}^\kappa)^\top (\mathbf{I}_2 + \nabla \boldsymbol{\chi}_{m,k'}^\kappa) \right\rangle. \quad (3.32)$$

This is the purpose of the next lemma.

**Lemma 3.2.** *Assume that  $\kappa$  satisfies (3.27). Then exists a constant  $C(\beta) < \infty$  such that*

$$|\mathbf{J}_m^\kappa(t) - \mathbf{E}_m^\kappa(t)| \leq \frac{C a_m^2 \varepsilon_m^4}{\kappa} \left( \frac{\varepsilon_m^2}{\kappa \tau_m} \right). \quad (3.33)$$

*Proof.* Observe that

$$\mathbf{E}_m^\kappa(t) = \sum_{k,k' \in 2\mathbb{Z}+1} \xi_{m,k}(t) \xi_{m,k'}(t) \kappa \left( \mathbf{I}_2 + \langle (\nabla \boldsymbol{\chi}_{m,k}^\kappa)^\top \nabla \boldsymbol{\chi}_{m,k'}^\kappa \rangle \right).$$

The supports of  $\xi_{m,k}$  and  $\xi_{m,k'}$  have nonempty intersection only if  $k, k' \in 2\mathbb{Z}+1$  satisfy  $|k - k'| \leq 2$ . On the other hand, we see from the formulas (3.9) and (3.5) that

$$|k - k'| = 2 \implies (\nabla \boldsymbol{\chi}_{m,k}^\kappa)^\top(t, x) \nabla \boldsymbol{\chi}_{m,k'}^\kappa(t, x') = 0, \quad \forall x, x' \in \mathbb{T}^2, \quad t \in \mathbb{R}. \quad (3.34)$$

Therefore, the only pairs  $k, k'$  contributing to the sum satisfy  $k = k'$ . We deduce that

$$\mathbf{E}_m^\kappa(t) = \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k}(t)^2 \kappa \left( \mathbf{I}_2 + \langle (\nabla \boldsymbol{\chi}_{m,k}^\kappa)^\top \nabla \boldsymbol{\chi}_{m,k}^\kappa \rangle \right). \quad (3.35)$$

We next compute

$$\langle (\nabla \boldsymbol{\chi}_{m,k}^\kappa)^\top \nabla \boldsymbol{\chi}_{m,k}^\kappa \rangle = \langle (\nabla \mathbf{u}_{m,k})^t \nabla \mathbf{u}_{m,k} \rangle \left( \int_{-\infty}^t \hat{\zeta}_{m,l_k}(s) \zeta_{m,k}(s) \exp\left(\frac{4\pi^2 \kappa}{\varepsilon_m^2}(s-t)\right) ds \right)^2.$$

We see from (3.5) that

$$((\nabla \mathbf{u}_{m,k})^t \nabla \mathbf{u}_{m,k})(x) = \begin{cases} 16\pi^4 a_m^2 \sin^2\left(\frac{2\pi x_1}{\varepsilon_m}\right) \mathbf{e}_2 \otimes \mathbf{e}_2 & \text{if } k \in 4\mathbb{Z}+1, \\ 16\pi^4 a_m^2 \sin^2\left(\frac{2\pi x_2}{\varepsilon_m}\right) \mathbf{e}_1 \otimes \mathbf{e}_1 & \text{if } k \in 4\mathbb{Z}+3, \\ 0 & \text{if } k \in 2\mathbb{Z}. \end{cases} \quad (3.36)$$

and thus

$$\langle (\nabla \mathbf{u}_{m,k})^t \nabla \mathbf{u}_{m,k} \rangle = 8\pi^4 a_m^2 \cdot \begin{cases} \mathbf{e}_2 \otimes \mathbf{e}_2 & \text{if } k \in 4\mathbb{Z} + 1, \\ \mathbf{e}_1 \otimes \mathbf{e}_1 & \text{if } k \in 4\mathbb{Z} + 3, \\ 0 & \text{if } k \in 2\mathbb{Z}. \end{cases} \quad (3.37)$$

Using (2.25) and (2.29), we see that

$$\left| \hat{\zeta}_{m,l_k}(t) \zeta_{m,k}(t) - \frac{4\pi^2 \kappa}{\varepsilon_m^2} \int_{-\infty}^t \hat{\zeta}_{m,l_k}(s) \zeta_{m,k}(s) \exp\left(\frac{4\pi^2 \kappa}{\varepsilon_m^2}(s-t)\right) ds \right| \leq C \left( \frac{C \varepsilon_m^2}{\kappa \tau_m} \right). \quad (3.38)$$

Combining the above and comparing to the formula for  $\mathbf{J}_m^\kappa$  in (3.18) yields (3.33).  $\square$

**Lemma 3.3.** *There exists  $C(\beta) \in [1, \infty)$  such that, for every  $\kappa > 0$  and  $m \in \mathbb{N}$ ,*

$$\left| \bar{\mathbf{K}}_m^\kappa - \left( \kappa + \frac{9a_m^2 \varepsilon_m^4}{80\kappa} \right) \mathbf{I}_2 \right| \leq \frac{a_m^2 \varepsilon_m^4}{\kappa} \left( \frac{C \varepsilon_m^2}{\kappa \tau_m} + C \varepsilon_{m-1}^\delta \right). \quad (3.39)$$

*Proof.* Starting from (3.18), we find that

$$\bar{\mathbf{K}}_m^\kappa - \kappa \mathbf{I}_2 = \frac{\pi^2 a_m^2 \varepsilon_m^2}{\tau_m''} \sum_{k \in 2\mathbb{Z}+1} \int_{-\frac{1}{2}\tau_m''}^{\frac{1}{2}\tau_m''} \hat{\zeta}_{m,l_k}(t) \zeta_{m,k}(t) \int_{-\infty}^t \hat{\zeta}_{m,l_k}(s) \zeta_{m,k}(s) \exp\left(\frac{4\pi^2 \kappa}{\varepsilon_m^2}(s-t)\right) ds dt \mathbf{I}_2.$$

It therefore suffices to show that

$$\begin{aligned} \left| \frac{9\varepsilon_m^2}{80\pi^2 \kappa} - \sum_{k \in 2\mathbb{Z}+1} \int_{-\frac{1}{2}\tau_m''}^{\frac{1}{2}\tau_m''} \int_{-\infty}^t \hat{\zeta}_{m,l_k}(t) \hat{\zeta}_{m,l_k}(s) \zeta_{m,k}(t) \zeta_{m,k}(s) \exp\left(\frac{4\pi^2 \kappa}{\varepsilon_m^2}(s-t)\right) ds dt \right| \\ \leq \frac{\varepsilon_m^2}{\kappa} \left( \frac{\varepsilon_m^2}{\kappa \tau_m} + C \varepsilon_{m-1}^\delta \right). \end{aligned} \quad (3.40)$$

Using (2.22), we have that

$$\begin{aligned} \int_{-\infty}^t |\zeta_{m,k}(t) - \zeta_{m,k}(s)| \exp\left(\frac{4\pi^2 \kappa}{\varepsilon_m^2}(s-t)\right) ds &\leq \|\partial_t \zeta_{m,k}\|_{L^\infty(\mathbb{R})} \int_{-\infty}^t (t-s) \exp\left(\frac{4\pi^2 \kappa}{\varepsilon_m^2}(s-t)\right) ds \\ &\leq C \tau_m^{-1} \int_{-\infty}^t (t-s) \exp\left(\frac{4\pi^2 \kappa}{\varepsilon_m^2}(s-t)\right) ds = \frac{C \varepsilon_m^4}{\kappa^2 \tau_m}. \end{aligned}$$

On the other hand,

$$\begin{aligned} \left| \sum_{k \in 2\mathbb{Z}+1} \int_{-\frac{1}{2}\tau_m''}^{\frac{1}{2}\tau_m''} \hat{\zeta}_{m,l_k}(t) \zeta_{m,k}(t)^2 \int_{-\infty}^t \hat{\zeta}_{m,l_k}(s) \exp\left(\frac{4\pi^2 \kappa}{\varepsilon_m^2}(s-t)\right) ds dt - \frac{\varepsilon_m^2}{4\pi^2 \kappa} \cdot \underbrace{\frac{1}{2} \int_{-\frac{1}{2}\tau_m}^{\frac{1}{2}\tau_m} \zeta_{m,0}^2(t) dt}_{= \frac{9}{10} \text{ by (2.23)}} \right| \\ \leq \frac{C \varepsilon_m^2 \tau_m'}{\kappa \tau_m''} \leq \frac{C \varepsilon_m^2 \varepsilon_{m-1}^\delta}{\kappa}. \end{aligned}$$

The triangle inequality and the previous two displays yield (3.40). The proof is now complete.  $\square$

As previously mentioned, we will apply (3.11) and (3.39) when the factor  $\varepsilon_m^2/(\kappa \tau_m)$  on the right side of (3.39) is very small, typically a small positive power of  $\varepsilon_m$ : see (3.46) below. Therefore, loosely sense, we have that

$$\bar{\mathbf{K}}_m^\kappa \approx \kappa + \frac{9a_m^2 \varepsilon_m^4}{80\kappa}. \quad (3.41)$$

We now define a sequence  $\{\kappa_m\}$  of *renormalized diffusivities*, starting from a given “molecular” diffusivity  $\kappa$ , by the recursion

$$\begin{cases} \kappa_{m-1} = \bar{\mathbf{K}}_m^{\kappa_m} & m \in \{1, \dots, M\}, \\ \kappa_M = \kappa. \end{cases} \quad (3.42)$$

The idea is that  $\kappa_{m-1}$  represents an effective (or “eddy”) diffusivity observed at scale  $\varepsilon_m$ , from the cumulative effects of the diffusion term  $\kappa\Delta$  and all the oscillations in the vector field  $\mathbf{b}$  with wavelengths smaller than  $\varepsilon_{m-1}$ . We imagine that we have homogenized all scales below that of  $\varepsilon_{m-1}$  and witnessed an enhancement of diffusivity which results in an effective diffusivity of  $\kappa_{m-1}$ . We choose the initial scale  $M$  in such a way that  $\varepsilon_M$  is the critical scale at which the vector field  $\mathbf{b}$  and the diffusion interact in such a way that the recursion (3.42) stays under control.

The next lemma states that, for certain particular values of the molecular diffusivity  $\kappa$ , we can control the entire sequence of renormalized diffusivities. We denote by  $\mathcal{K}$  the set of *permissible diffusivities*, defined by

$$\mathcal{K} := \bigcup_{m=1}^{\infty} \left[ \frac{1}{2} \varepsilon_m^{\frac{2\beta}{q+1}}, 2 \varepsilon_m^{\frac{2\beta}{q+1}} \right]. \quad (3.43)$$

Recall that  $q$  is introduced in above in (2.2) and is also related to  $\beta$  by the formula (2.3). The definitions of the exponents  $\delta$  and  $\gamma$  appearing in the lemma statement below are also given above, in (2.5) and (2.7), respectively.

The reason that we can only control sequences  $\{\kappa_m\}$  satisfying the satisfying recurrence (3.42) with initial values  $\kappa$  belonging to this set  $\mathcal{K}$  has to do with the instability of the recurrence formula in (3.41)–(3.42). This does not appear to be a technical artifact of our proof, but rather a property of the recursion itself.

**Lemma 3.4** (Control of the renormalized diffusivities). *Suppose  $\kappa \in \mathcal{K}$  and let  $M \in \mathbb{N}$  be such that*

$$\frac{1}{2} \varepsilon_M^{\frac{2\beta}{q+1}} \leq \kappa \leq 2 \varepsilon_M^{\frac{2\beta}{q+1}}. \quad (3.44)$$

*Define a finite sequence  $\kappa_M, \kappa_{M-1}, \dots, \kappa_0$  by the recurrence (3.42), starting from  $\kappa_M := \kappa$ . Then there exist universal constants  $0 < c < C < \infty$  such that, for every  $m \in \{0, \dots, M-1\}$ ,*

$$c a_m \varepsilon_m^{2+\gamma} \leq \kappa_m \leq C a_m \varepsilon_m^{2+\gamma} \quad (3.45)$$

and

$$c \varepsilon_{m-1}^{2\delta} \leq \frac{\varepsilon_m^2}{\kappa_m \tau_m} \leq C \varepsilon_{m-1}^{2\delta}. \quad (3.46)$$

*Proof.* We proceed by first establishing (3.45) for the a different sequence  $\{\kappa'_m\}$ , defined by

$$\begin{cases} \kappa'_{m-1} = \kappa'_m + \frac{9a_m^2 \varepsilon_m^4}{80\kappa'_m}, & m \in \{1, \dots, M\}, \\ \kappa'_M = \kappa. \end{cases} \quad (3.47)$$

According to our (imprecise) shorthand (3.41), we have reasons to expect that  $\kappa'_m$  is close to  $\kappa_m$ . Once we have proved (3.45), we will argue that the two sequences are indeed close enough that we may obtain essentially the same estimate for  $\kappa_m$ . Recall that the parameter  $\gamma$  defined in (2.7) satisfies, in view of (2.11) and (2.4),

$$a_m \varepsilon_m^{2+\gamma} = \varepsilon_m^{\beta+\gamma}. \quad (3.48)$$

*Step 1.* We prove that there exist universal constants  $0 < c \leq C < \infty$  such that

$$ca_m \varepsilon_m^{2+\gamma} \leq \kappa'_m \leq Ca_m \varepsilon_m^{2+\gamma}, \quad \forall m \in \{0, \dots, M-1\}. \quad (3.49)$$

Denote

$$s_m := \frac{\kappa'_m \sqrt{80/9}}{a_m \varepsilon_m^{2+\gamma}}. \quad (3.50)$$

We may rewrite the recurrence in (3.47) in terms of  $s_m$  as

$$s_{m-1} = \left( \frac{\varepsilon_m}{\varepsilon_{m-1}} \right)^\beta \varepsilon_m^{-\gamma} \varepsilon_{m-1}^{-\gamma} \cdot s_m \left( \varepsilon_m^{2\gamma} + \frac{1}{s_m^2} \right). \quad (3.51)$$

Notice that the exponent  $\gamma$  has been chosen so that it satisfies  $q(\beta - \gamma) = \beta + \gamma$ . Hence by (2.10) we have

$$\left| \left( \frac{\varepsilon_m}{\varepsilon_{m-1}} \right)^\beta \cdot \varepsilon_m^{-\gamma} \varepsilon_{m-1}^{-\gamma} - 1 \right| = \left| \left( \frac{\varepsilon_m}{\varepsilon_{m-1}^q} \right)^\beta - 1 \right| \leq \frac{C \varepsilon_{m-1}}{\beta} \leq C \varepsilon_{m-1}. \quad (3.52)$$

We therefore obtain from (3.51) that

$$\left| s_{m-1} - s_m \left( \varepsilon_m^{2\gamma} + \frac{1}{s_m^2} \right) \right| \leq C \varepsilon_{m-1} s_m \left( \varepsilon_m^{2\gamma} + \frac{1}{s_m^2} \right). \quad (3.53)$$

In view of (3.48), the condition (3.44) can be written as

$$\frac{1}{2} \varepsilon_M^{-2\gamma} \leq s_M \leq 2 \varepsilon_M^{-2\gamma}.$$

We deduce from this and (3.53) that

$$s_{M-1} \leq 2 + 2 \varepsilon_M^{2\gamma} + C \varepsilon_{M-1} \leq 4$$

and

$$s_{M-1} \geq \frac{1}{2} - C \varepsilon_{M-1} \geq \frac{1}{4}.$$

Similarly, it is easy to check that

$$\max \left\{ s_{m-1}, \frac{1}{s_{m-1}} \right\} \leq \max \left\{ s_m, \frac{1}{s_m} \right\} \left( 1 + C \varepsilon_{m-1}^{2q\theta \wedge 1} \right).$$

An iteration of the latter inequality therefore yields

$$\max \left\{ s_{m-1}, \frac{1}{s_{m-1}} \right\} \leq 4 \prod_{j=m}^M \left( 1 + C \varepsilon_{j-1}^{2q\theta \wedge 1} \right) \leq 4 \left( 1 + C \varepsilon_{m-1}^{2q\theta \wedge 1} \right) \leq C. \quad (3.54)$$

The proof of (3.49) is now complete.

*Step 2.* We show that (3.49) implies (3.45). By (2.11), (2.12) and (3.49), we observe that

$$\frac{\varepsilon_m^2}{\kappa'_m \tau_m} \simeq \varepsilon_m^{-\gamma} a_m^{-1} \tau_m^{-1} \simeq \varepsilon_{m-1}^{(2-\beta)(q-1)-q\gamma-2\delta}.$$

By the definitions of the exponents in (2.5) and (2.7), we have that

$$(2-\beta)(q-1)-q\gamma = (q-1) \left( 2 - \frac{2q+1}{q+1} \beta \right) = 4\delta.$$

Hence

$$c\varepsilon_{m-1}^{2\delta} \leq \frac{\varepsilon_m^2}{\kappa'_m \tau_m} \leq C\varepsilon_{m-1}^{2\delta}. \quad (3.55)$$

Arguing by induction, suppose that for some  $n \in \{1, \dots, M-1\}$ , we have

$$\frac{1}{2} \leq \frac{\kappa'_m}{\kappa_m} \leq 2, \quad \forall m \in \{n, \dots, M-1\}. \quad (3.56)$$

Then, using also (3.49), we deduce that, for every  $m \in \{n, \dots, M-1\}$ ,

$$(1 - C\varepsilon_m^{2\gamma}) \left( \frac{\kappa_m}{\kappa'_m} \right)^2 \leq \frac{1 + \frac{9a_m^2 \varepsilon_n^4}{80(\kappa'_m)^2}}{1 + \frac{9a_m^2 \varepsilon_n^4}{80\kappa_m^2}} \leq (1 + C\varepsilon_m^{2\gamma}) \left( \frac{\kappa_m}{\kappa'_m} \right)^2.$$

Next we use (3.39), (3.55) and (3.56) to get

$$(1 - C\varepsilon_{n-1}^{2\delta}) \left( 1 + \frac{9a_n^2 \varepsilon_n^4}{80\kappa_n^2} \right) \leq \frac{\kappa_{n-1}}{\kappa_n} \leq (1 + C\varepsilon_{n-1}^{2\delta}) \left( 1 + \frac{9a_n^2 \varepsilon_n^4}{80\kappa_n^2} \right) \quad (3.57)$$

Putting these together and using the exact recursion formula for  $\kappa'_{n-1}$ , we get that

$$\max \left\{ \frac{\kappa_{n-1}}{\kappa'_{n-1}}, \frac{\kappa'_{n-1}}{\kappa_{n-1}} \right\} \leq (1 + C\varepsilon_{n-1}^{2(\delta \wedge \gamma)}) \frac{\kappa_n}{\kappa'_n}.$$

Iterating this and using  $\kappa_M = \kappa'_N$ , we find that

$$\max \left\{ \frac{\kappa_{n-1}}{\kappa'_{n-1}}, \frac{\kappa'_{n-1}}{\kappa_{n-1}} \right\} \leq \prod_{j=n}^N (1 + C\varepsilon_{j-1}^{2(\delta \wedge \gamma)}) \leq (1 + C\varepsilon_{n-1}^{2(\delta \wedge \gamma)}). \quad (3.58)$$

This allows us to remove the condition (3.56) and replace it with  $C\varepsilon_{n-1}^{2(\delta \wedge \gamma)} \leq 1$ ; that is, for some  $n_0(\text{data}) \in \mathbb{N}$ , we have that the inequality (3.58) holds for every  $n \geq n_0$ . However, for  $m \leq n_0$ , we have  $c \leq \min\{\kappa_m, \kappa'_m\} \leq \max\{\kappa_m, \kappa'_m\} \leq C$ , and so we have shown that

$$\max_{m \in \{0, \dots, M-1\}} \max \left\{ \frac{\kappa_m}{\kappa'_m}, \frac{\kappa'_m}{\kappa_m} \right\} \leq C. \quad (3.59)$$

In view of (3.49) and (3.55), the proof of the lemma is now complete.  $\square$

In most of the rest of the paper, we assume that the molecular diffusivity constant  $\kappa$  belongs to the set  $\mathcal{K}$  of permissible diffusivities defined in (3.43), so that the bounds of Lemma 3.4 are valid. Incidentally, the reason we are only able to obtain anomalous diffusion along a subsequence of  $\kappa$ 's in Theorem 1.1 is due to the restriction in Lemma 3.4.

#### 4. The multiscale ansatz

Now that we have constructed the vector field  $\mathbf{b}$  and defined the renormalized diffusivities, we are ready to begin the proof of anomalous diffusion. This will require some delicate asymptotic expansions which will take us the next several sections to develop.

Throughout, we fix a molecular diffusivity  $\kappa \in \mathcal{K}$ , with the set  $\mathcal{K}$  defined in (3.43). We let  $M$  denote the positive integer satisfying (3.44), and we let the finite sequence  $\kappa_M, \kappa_{M-1}, \dots, \kappa_0$  be

defined by (3.42). We also select an initial datum, which is  $\mathbb{Z}^2$ -periodic function  $\theta_0 \in C^\infty(\mathbb{T}^2)$  with zero mean,

$$\langle \theta_0 \rangle = \int_{\mathbb{T}^2} \theta_0(x) dx = 0, \quad (4.1)$$

and which satisfies the quantitative analyticity condition (recall the notation in (1.22))

$$\|\nabla^n \theta_0\|_{L^2(\mathbb{T}^2)} \leq \|\theta_0\|_{L^2(\mathbb{T}^2)} n! R_{\theta_0}^{-n}, \quad \forall n \in \mathbb{N}. \quad (4.2)$$

For each  $m \in \{0, \dots, M\}$ , we let  $\theta_m$  denote the solution of the initial-value problem

$$\begin{cases} \partial_t \theta_m - \kappa_m \Delta \theta_m + \mathbf{b}_m \cdot \nabla \theta_m = 0 & \text{in } (0, \infty) \times \mathbb{R}^2, \\ \theta_m = \theta_0 & \text{on } \{0\} \times \mathbb{R}^2. \end{cases} \quad (4.3)$$

Recall that  $\mathbf{b}_m$  is defined in (2.36). In other words,  $\theta_m$  is the solution of the modified equation in which the stream function  $\phi$  has been replaced by  $\phi_m$ , essentially removing the oscillations of  $\phi$  with wavelengths smaller than  $\varepsilon_m$ . Since  $\phi_m$  is smooth, the equation can be written in terms of the vector field  $\mathbf{b}_m$ , as above, but it is often more convenient to write it in terms of the stream function  $\phi_m$  as

$$\partial_t \theta_m - \nabla \cdot (\kappa_m \mathbf{I}_2 + \phi_m \sigma) \nabla \theta_m = 0 \quad \text{in } (0, \infty) \times \mathbb{R}^2. \quad (4.4)$$

It is clear that  $\theta_m \in C^\infty((0, \infty) \times \mathbb{R}^2)$  and, for each time  $t \in (0, \infty)$ , the function  $\theta_m(t, \cdot)$  has zero mean and is  $\mathbb{Z}^2$ -periodic. Note that, (4.3) in the case  $m = 0$ , extends the domain of the given function  $\theta_0$  from  $\mathbb{T}^2$ , which we identify with  $\{0\} \times \mathbb{T}^2$ , to  $[0, \infty) \times \mathbb{T}^2$ .

In order to prove Theorem 1.1, we will propagate lower bounds on the energy dissipation of the solutions of (4.3) from  $m - 1$  to  $m$ . In fact, the key step is show that, for every  $m \in \{0, \dots, M\}$  with  $\varepsilon_{m-1} \leq R_{\theta_0}$ , we have

$$\left| \frac{\kappa_m \|\nabla \theta_m\|_{L^2((0,1) \times \mathbb{T}^2)}^2}{\kappa_{m-1} \|\nabla \theta_{m-1}\|_{L^2((0,1) \times \mathbb{T}^2)}^2} - 1 \right| \leq C \varepsilon_{m-1}^\delta. \quad (4.5)$$

This estimate is proved in Proposition 5.2, below. From (4.5), it is a simple matter to obtain the lower bound on the energy dissipation in Theorem 1.1, as we will see.

The proof of (4.5) is based on the informal idea that the equation for  $\theta_m$  should homogenize to the equation for  $\theta_{m-1}$ . To see why we should expect this, write the equation for  $\theta_m$  as

$$(\partial_t + \mathbf{b}_{m-1} \cdot \nabla) \theta_m - \nabla \cdot (\kappa_m \mathbf{I}_2 + \tilde{\psi}_m \sigma) \nabla \theta_m = 0 \quad \text{in } (0, \infty) \times \mathbb{R}^2, \quad (4.6)$$

where we define

$$\tilde{\psi}_m(t, x) := \phi_m(t, x) - \phi_{m-1}(t, x) = \sum_{k \in 2\mathbb{Z}+1} \hat{\zeta}_{m, l_k}(t) \zeta_{m, k}(t) \psi_{m, k}(X_{m-1, l_k}^{-1}(t, x)).$$

We view the vector field  $\mathbf{b}_{m-1}$  in the transport term in the left of (4.6) as *slow*, as well as the corresponding flows  $X_{m-1, l}$  and inverse flows  $X_{m-1, l}^{-1}$ . In contrast, we consider the coefficient matrix  $\kappa_m \mathbf{I}_2 + \tilde{\psi}_m \sigma$  in the second-order part of the operator to be *fast*. Moreover, if we change variables to Lagrangian coordinates with respect to the “slow” flows, then the transport operator  $\partial_t + \mathbf{b}_{m-1} \cdot \nabla$  becomes simply  $\partial_t$  and diffusion operator  $\nabla \cdot (\kappa_m \mathbf{I}_2 + \tilde{\psi}_m \sigma) \nabla$  becomes  $\nabla \cdot (\kappa_m \mathbf{I}_2 + \psi_m \sigma) \nabla$ , which is a shear flow which switches between the horizontal and vertical directions. Given the discussion in

the previous section, we expect the fast diffusive operator to homogenize to  $\kappa_{m-1}\Delta$ , which leaves us with the equation for  $\theta_{m-1}$  in the original coordinates.

When we speak here of “homogenization” we do not intend for the reader to understand this too literally: no limit is taken, rather the equations will be shown to be close in quantitative sense which is small relative to a power of  $\varepsilon_{m-1}$ .

To make this idea precise, we introduce an multiscale ansatz for  $\theta_m$ , denoted by  $\tilde{\theta}_m$ , which is built from  $\theta_{m-1}$  and the correctors defined in the previous section. Our strategy is very simple: we will plug  $\tilde{\theta}_m$  into the equation for  $\theta_m$  and estimate the error. We will show that it is small enough to conclude that  $\tilde{\theta}_m$  is close to  $\theta_m$ . Since we built  $\tilde{\theta}_m$  from  $\theta_{m-1}$ , we will be able to relate  $\theta_m$  to  $\theta_{m-1}$  and, in particular, obtain (4.5).

The definition of  $\tilde{\theta}_m$  is motivated by the usual two-scale ansatz used in classical homogenization, in which one attaches the periodic correctors to the (usually smooth) solution of the macroscopic equation, as explained in the discussion above (1.19). However, it is necessarily more complicated than the naive guess (1.19), for several reasons.

First of all, there are actually three different “fast” scales in the equation for  $\theta_m$ :

- (i) the smallest spatial scale  $\varepsilon_m$ , which is the length scale of the shear flows;
- (ii) the time scale  $\tau_m$ , on which the shear flows switch directions;
- (iii) the time scale  $\tau_m''$ , on which the Lagrangian flows  $X_{m-1}$  must refresh.

We should think of these three fast scales as being well-separated, with the spatial scale  $\varepsilon_m$  being the smallest/fastest. This means that each of these three scales must be separately homogenized! Since homogenization estimates require smoothness of the macroscopic data, we must be very careful to maintain sufficient regularity estimates when we homogenize the time scales. This is the reason we spend so much effort proving estimates on  $T_{m-1}$  and  $\tilde{H}_m$  later in this section.

A second complication is due to the need to compose with the Lagrangian flows. As we have seen informally above, when we homogenize the spatial oscillations (the shear flows), we need to work in Lagrangian coordinates. Rather than actually switching our coordinate system, our definition of  $\tilde{\theta}_m$  will involve compositions with the inverse flows  $X_{m-1}^{-1}$ . Unfortunately, the distortion caused by these flows cannot be ignored, and we must therefore introduce corrections in the equation for  $\theta_{m-1}$ . This is the reason for the appearance of the matrix  $\mathbf{s}_{m-1}$  defined in (4.7), and its role in the definition of  $T_{m-1}$ .

We select  $m \in \{1, \dots, M\}$  which is fixed throughout the rest of this section. We also employ the following two notational conventions, which are in force throughout most of the rest of the paper:

- We use the correctors  $\chi_m^\kappa$  and matrices  $\mathbf{J}_m^\kappa$ ,  $\mathbf{K}_m^\kappa$  and  $\bar{\mathbf{K}}_m^\kappa$  introduced in the previous section with  $\kappa = \kappa_m$  (and never any other choice of the parameter  $\kappa$ ). In order to lighten the notation, we drop the display of the dependence on  $\kappa_m$  from the superscripts, writing for example  $\chi_m$  and  $\mathbf{J}_m$  instead of  $\chi_m^{\kappa_m}$  and  $\mathbf{J}_m^{\kappa_m}$ . Recall that  $\bar{\mathbf{K}}_m^{\kappa_m} = \kappa_{m-1}$  by (3.42).
- We employ the convention that all function compositions are assumed to occur in the spatial variable only. In other words, as all function compositions involve the flows  $X_{m-1,l}$  and their inverses  $X_{m-1,l}^{-1}$  (see (2.37)), instead of writing  $(F(t, \cdot) \circ X_{m-1,l}(t, \cdot))(x)$  we will just write  $F \circ X_{m-1,l}$ .

In the next subsection, we will introduce the objects  $\mathbf{s}_{m-1}$ ,  $T_{m-1}$  and  $\tilde{H}_m$  that are needed in the definition of  $\tilde{\theta}_m$ , which is then given in Section 4.2. In Sections 4.3 and 4.4, we will prove important

regularity estimates on  $T_{m-1}$  and  $\tilde{H}_m$  which are needed in the following section. We will not see why our ansatz is a good one until the analysis in Section 5, which is where we plug it into the equation for  $\theta_m$  and compute the error. The definitions here are motivated by the computations in Section 5, and so we ask for the reader's patience if they seem a bit mysterious at first glance.

#### 4.1 Ingredients for the multiscale ansatz

We proceed by introducing equations with gradually more and more scales, starting from the largest scales, until we arrive at a guess for what  $\theta_m$  should look like.

We first introduce an equation with time oscillations on scale  $\tau_m''$ , which are due to the resetting of the flows. We define the matrix

$$\mathbf{s}_{m-1} := \mathbf{K}_m \sum_{l \in \mathbb{Z}} \hat{\xi}_{m,l} (\nabla X_{m-1,l} \circ X_{m-1,l}^{-1} - \mathbf{I}_2). \quad (4.7)$$

At this point, we want to modify the equation for  $\theta_{m-1}$  by introducing a diffusion coefficient which oscillates on the time scale  $\tau_m''$ . We will call the resulting solution  $T_{m-1}$ .

The rough idea is to define  $T_{m-1} \in C^\infty([0, 1] \times \mathbb{T}^2)$  to be the solution of the initial-value problem

$$\begin{cases} \partial_t T_{m-1} - \nabla \cdot (\mathbf{K}_m + \mathbf{s}_{m-1}) \nabla T_{m-1} + \mathbf{b}_{m-1} \cdot \nabla T_{m-1} = 0 & \text{in } (0, \infty) \times \mathbb{R}^2, \\ T_{m-1} = \theta_0 & \text{on } \{0\} \times \mathbb{R}^2. \end{cases} \quad (4.8)$$

It is not difficult to see why we should expect this equation to homogenize to the one for  $\theta_{m-1}$ : the principal part of the diffusion matrix in (4.8) is  $\mathbf{K}_m$ , which has periodic oscillations in time only, with period  $\tau_m''$  and a mean which very close to  $\bar{\mathbf{K}}_m = \kappa_{m-1}$ . The matrix  $\mathbf{s}_{m-1}$  is lower-order compared to  $\mathbf{K}_m$ , due to (2.67). The reason for including it in the equation of  $T_{m-1}$  has to do with the need to anticipate some errors arising in the analysis of the smaller (spatial) scales, due to the change to Lagrangian coordinates. Note that we may also write the equation (4.8) as

$$\partial_t T_{m-1} - \nabla \cdot (\mathbf{K}_m + \mathbf{s}_{m-1} + \phi_{m-1} \sigma) \nabla T_{m-1} = 0 \quad \text{in } (0, \infty) \times \mathbb{R}^2. \quad (4.9)$$

We do *not* actually define  $T_{m-1}$  to be the solution of (4.8). We will instead define it as an approximate solution of (4.8) through an iteration procedure. The advantage of this is that it allows us to prove better regularity estimates for  $T_{m-1}$ . Indeed, the best lower bound on the matrix  $\mathbf{K}_m$  is  $\kappa_m \mathbf{I}_2$ , which is much less than  $\kappa_{m-1} \mathbf{I}_2$ , even though  $\mathbf{K}_m$  is larger than the latter on a proportion of times at least  $1 - C\varepsilon_m^\delta$ . Nevertheless, if this is used in the energy estimates we will get very pessimistic regularity bounds on  $T_{m-1}$  compared to those we have for  $\theta_{m-1}$  in Lemma 4.1 below. To get better bounds, we work with an approximate solution of (4.8) which is constructed as follows.

We will choose a large number, which represents the number of iteration steps in our definition of  $T_{m-1}$ , and for convenience we may take the large integer  $N_*$  defined in (2.6). We initialize the iteration by setting

$$T_{m-1}^{(0)} := \theta_{m-1}. \quad (4.10)$$

For every  $1 \leq i \leq N_*$ , we recursively define  $T_{m-1}^{(i)}$  to be the solution of the initial-value problem

$$\begin{cases} \partial_t T_{m-1}^{(i)} - \kappa_{m-1} \Delta T_{m-1}^{(i)} + \mathbf{b}_{m-1} \cdot \nabla T_{m-1}^{(i)} = \nabla \cdot (\mathbf{K}_m - \kappa_{m-1} \mathbf{I}_2 + \mathbf{s}_{m-1}) \nabla T_{m-1}^{(i-1)} & \text{in } (0, \infty) \times \mathbb{R}^2, \\ T_{m-1}^{(i)} = \theta_0 & \text{on } \{0\} \times \mathbb{R}^2. \end{cases} \quad (4.11)$$

It is clear that  $T_{m-1}^{(i)} \in C^\infty([0, \infty) \times \mathbb{R}^2)$ . Finally, we define

$$T_{m-1} := T_{m-1}^{(N_*)}. \quad (4.12)$$

By construction,  $T_{m-1}$  satisfies

$$\begin{cases} \partial_t T_{m-1} - \nabla \cdot (\mathbf{K}_m + \mathbf{s}_{m-1}) \nabla T_{m-1} + \mathbf{b}_{m-1} \cdot \nabla T_{m-1} = \nabla \cdot \mathbf{e}_{m-1} & \text{in } (0, \infty) \times \mathbb{R}^2, \\ T_{m-1} = \theta_0 & \text{on } \{0\} \times \mathbb{R}^2, \end{cases} \quad (4.13)$$

where the error  $\mathbf{e}_{m-1}$  is given by

$$\mathbf{e}_{m-1} := (\mathbf{K}_m - \kappa_{m-1} \mathbf{I}_2 + \mathbf{s}_{m-1}) \nabla (T_{m-1}^{(N_*)-1} - T_{m-1}^{(N_*)}). \quad (4.14)$$

Comparing (4.13) with (4.8), we see that the “true”  $T_{m-1}$  makes an error  $\nabla \cdot \mathbf{e}_{m-1}$  in solving the advection-diffusion equation. We will however show that error  $\mathbf{e}_{m-1}$  will be very small, in fact it can be made “very small” since  $N_*$  is very large (see (2.6)).

Below in Section 4.3 we will show that the difference  $T_{m-1} - \theta_{m-1}$  is small (see Lemma 4.4 for the precise statement). This amounts to homogenizing the temporal oscillations due to the switching of the flows, which have period  $\tau_m''$ .

We would next like to write down an equation like (4.9), but with  $\mathbf{J}_m$  in place of  $\mathbf{K}_m$ . That is, we want to include the temporal oscillations due to switching between horizontal and vertical shear flows. Recall that the function  $\mathbf{J}_m$  is essentially the sum of products of  $\tau_m$ -periodic and  $\tau_m''$ -periodic functions of time only. When the faster time scale is averaged out of  $\mathbf{J}_m$ , one obtains  $\mathbf{K}_m$ , up to very small errors: see (3.20) (3.25), and (3.26), above. We could write this equation perhaps as

$$\begin{cases} \partial_t \tilde{T}_{m-1} - \nabla \cdot (\mathbf{J}_m + \mathbf{r}_{m-1}) \nabla \tilde{T}_{m-1} + \mathbf{b}_{m-1} \cdot \nabla \tilde{T}_{m-1} = 0 & \text{in } (0, \infty) \times \mathbb{R}^2, \\ \tilde{T}_{m-1} = \theta_0 & \text{on } \{0\} \times \mathbb{R}^2, \end{cases} \quad (4.15)$$

where  $\mathbf{r}_{m-1}$  is defined as in (4.7), with  $\mathbf{J}_m$  in place of  $\mathbf{K}_m$ . We will show the difference  $\tilde{T}_{m-1} - T_{m-1}$  is, up to errors we are able to neglect, given by an expansion which we now introduce.

We introduce a function  $\tilde{H}_m$  which is intended to represent, to leading order, the difference between  $\tilde{T}_{m-1} - T_{m-1}$ . The first idea is to take it to solve the transport-type equation<sup>11</sup>

$$\begin{cases} (\partial_t + \mathbf{b}_{m-1} \cdot \nabla) \tilde{H}_m = \tilde{G}_m & \text{in } (0, \infty) \times \mathbb{R}^2, \\ \tilde{H}_m = 0 & \text{on } \{0\} \times \mathbb{R}^2, \end{cases} \quad (4.16)$$

where<sup>12</sup>

$$\begin{aligned} \tilde{G}_m &:= (\mathbf{J}_m - \mathbf{K}_m) : \sum_{l \in \mathbb{Z}} \hat{\xi}_{m,l} \nabla (\nabla (T_{m-1} \circ X_{m-1,l}) \circ X_{m-1,l}^{-1}) \\ &= \nabla \cdot (\mathbf{J}_m - \mathbf{K}_m) \sum_{l \in \mathbb{Z}} \hat{\xi}_{m,l} (\nabla X_{m-1,l} \circ X_{m-1,l}^{-1}) \nabla T_{m-1}. \end{aligned} \quad (4.18)$$

<sup>11</sup>See (4.21) for the actual definition of  $\tilde{H}_m$ , and (4.22)–(4.23) for the equation it solves.

<sup>12</sup>Note that, in the second line of (4.18), we used our notational convention (introduced in (3.7)) of writing vector-valued functions as row vectors and gradients of scalars as column vectors. Hence

$$\nabla (T_{m-1} \circ X_{m-1,k}) \circ X_{m-1,k}^{-1} = (\nabla X_{m-1,k} \circ X_{m-1,k}^{-1}) \nabla T_{m-1}. \quad (4.17)$$

We get this by subtracting the equations for  $T_{m-1}$  and  $\tilde{T}_{m-1}$  and then ignoring the diffusion term  $\nabla \cdot (\mathbf{J}_m + \mathbf{r}_{m-1})(\nabla T_{m-1} - \nabla \tilde{T}_{m-1})$ . We will not however define  $\tilde{H}_m$  to be the solution of (4.16), because we are unable to prove sufficient regularity estimates for it. Instead we will define  $\tilde{H}_m$  explicitly in terms of known ingredients which approximate the solution of (4.16), but for which we can prove better regularity estimates.

In order to define  $\tilde{H}_m$ , we first define a function  $\tilde{H}_{m,r}$  as follows:

$$\tilde{H}_{m,r}(t, x) := \nabla \cdot \sum_{n=0}^{N_*-1} \mathbf{A}_{m,n,r}(t, x) \mathbf{q}_{m,n,r+1}^{\kappa_m}(t) \quad (4.19)$$

where  $\mathbf{q}_{m,n,r} = \mathbf{q}_{m,k,r}^{\kappa_m}$  as defined in (3.30) and the tensors  $\mathbf{A}_{m,n,r} = (\mathbf{A}_{m,n,r}^{ijk})_{i,j,k=1}^2$  are the 3-tensors recursively defined as follows:

$$\begin{cases} \mathbf{A}_{m,n,0}^{ijk} := -\delta_{ij} L_{m,n}^{\kappa_m}(t) \sum_{l \in \mathbb{Z}} \hat{\xi}_{m,l}(t) (\partial_k X_{m-1,l}^p \circ X_{m-1,l}^{-1}) \partial_p T_{m-1}, \\ \mathbf{A}_{m,n,r+1}^{ijk} := (\partial_t + \mathbf{b}_{m-1} \cdot \nabla) \mathbf{A}_{m,n,r}^{ijk} + \nabla_\ell \mathbf{b}_{m-1}^i \mathbf{A}_{m,n,r}^{\ell j k}. \end{cases} \quad (4.20)$$

In (4.19) we make the following convention regarding the contraction of indices: if  $\mathbf{q} = (\mathbf{q}^{jk})_{k,j=1}^2$  is a 2-tensor, and  $\mathbf{A} = (\mathbf{A}^{ijk})_{j,k,i=1}^2$  is a 3-tensor, then  $\nabla \cdot (\mathbf{A} \mathbf{q}) = \partial_i (\mathbf{A}^{ijk} \mathbf{q}^{jk})$ .

The definition in (4.20) above was made in view of the fact that

$$[(\partial_t + \mathbf{b}_{m-1} \cdot \nabla), \nabla] = \nabla \mathbf{b}_{m-1} \cdot \nabla$$

so that

$$\begin{aligned} (\partial_t + \mathbf{b}_{m-1} \cdot \nabla) \tilde{H}_{m,r} &= (\partial_t + \mathbf{b}_{m-1} \cdot \nabla) \nabla \cdot \sum_{n=0}^{N_*-1} \mathbf{A}_{m,n,r}(t, x) \mathbf{q}_{m,n,r+1}^{\kappa_m}(t) \\ &= \nabla \cdot \sum_{n=0}^{N_*-1} \mathbf{A}_{m,n,r+1} \mathbf{q}_{m,n,r+1}^{\kappa_m} - \nabla \cdot \sum_{n=0}^{N_*-1} \mathbf{A}_{m,n,r} \mathbf{q}_{m,n,r}^{\kappa_m} \end{aligned}$$

and therefore, by defining

$$\tilde{H}_m = \sum_{r=0}^{N_*/2} \tilde{H}_{m,r} \quad (4.21)$$

we obtain, by telescoping the resulting sum, that

$$(\partial_t + \mathbf{b}_{m-1} \cdot \nabla) \tilde{H}_m = -\nabla \cdot \sum_{n=0}^{N_*-1} \mathbf{A}_{m,n,0} \mathbf{q}_{m,n,0}^{\kappa_m} + \nabla \cdot \sum_{n=0}^{N_*-1} \mathbf{A}_{m,n,N_*/2} \mathbf{q}_{m,n,N_*/2}^{\kappa_m}.$$

We rewrite the first term as

$$\begin{aligned} &-\nabla \cdot \sum_{n=0}^{N_*-1} \mathbf{A}_{m,n,0}(t, x) \mathbf{q}_{m,n,0}^{\kappa_m}(t) \\ &= \partial_i \sum_{n=0}^{N_*-1} \left( \mathbf{j}_{m,n}^\kappa - \langle\langle \mathbf{j}_{m,n}^\kappa \rangle\rangle \right)_{jk} \left( \delta_{ij} L_{m,n}^{\kappa_m}(t) \sum_{l \in \mathbb{Z}} \hat{\xi}_{m,l}(t) \partial_k (T_{m-1} \circ X_{m-1,l}) \circ X_{m-1,l}^{-1} \right) \\ &= \nabla \cdot (\hat{\mathbf{J}}_m - \mathbf{K}_m) \sum_{l \in \mathbb{Z}} \hat{\xi}_{m,l} (\nabla X_{m-1,l} \circ X_{m-1,l}^{-1}) \nabla T_{m-1} \\ &= \tilde{G}_m + \nabla \cdot (\hat{\mathbf{J}}_m - \mathbf{J}_m) \sum_{l \in \mathbb{Z}} \hat{\xi}_{m,l} (\nabla X_{m-1,l} \circ X_{m-1,l}^{-1}) \nabla T_{m-1}. \end{aligned}$$

We therefore obtain that

$$(\partial_t + \mathbf{b}_{m-1} \cdot \nabla) \tilde{H}_m = \tilde{G}_m + \nabla \cdot \mathbf{d}_m \quad (4.22)$$

where we define  $\mathbf{d}_m$  by

$$\mathbf{d}_m := (\hat{\mathbf{J}}_m - \mathbf{J}_m) \sum_{l \in \mathbb{Z}} \hat{\xi}_{m,l} (\nabla X_{m-1,l} \circ X_{m-1,l}^{-1}) \nabla T_{m-1} + \sum_{n=0}^{N_*-1} \mathbf{A}_{m,n,N_*/2} \mathbf{q}_{m,n,N_*/2}^{\kappa_m}. \quad (4.23)$$

The  $\mathbf{d}_m$  error will be very small, proportional to  $\varepsilon_{m-1}^{\delta N_*}$  (which will be much less than  $\varepsilon_{m-1}^{1000}$ ). This is because of the closeness of  $\hat{\mathbf{J}}$  to  $\mathbf{J}$  in (3.25), and the fact that  $\mathbf{A}_{m,n,N_*}$  is similarly small, as we will show in Section 4.4.

As we will show in Lemma 4.7 below,  $\tilde{H}_m$  and its gradient are relatively small. In fact, its gradient is small enough that the error made by plugging it into the diffusion part of the operator is small and can be neglected. This says implicitly that the equation (4.15) homogenizes to (4.8), which takes care of the temporal oscillations on scale  $\tau_m$ .

## 4.2 Definition of the multiscale ansatz $\tilde{\theta}_m$

Without further ado, we can now present the two-scale ansatz  $\tilde{\theta}_m$ , which is defined by

$$\tilde{\theta}_m := T_{m-1} + \sum_{k,l \in \mathbb{Z}} \hat{\xi}_{m,l} \xi_{m,k} \tilde{\chi}_{m,k} (\nabla (T_{m-1} \circ X_{m-1,l}) \circ X_{m-1,l}^{-1}) + \tilde{H}_m \quad (4.24)$$

$$= T_{m-1} + \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} \tilde{\chi}_{m,k} (\nabla (T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1}) + \tilde{H}_m, \quad (4.25)$$

where  $\tilde{\chi}_{m,k}$  is the ‘‘twisted corrector’’ defined by

$$\tilde{\chi}_{m,k} := \chi_{m,k} \circ X_{m-1,k}^{-1}. \quad (4.26)$$

Recall that  $l_k$  is defined in (2.17) and satisfies (2.32). The second line (4.25) in the display above is valid due to (2.30) the fact that  $\chi_{m,k}$  vanishes if  $k$  is even and

$$\hat{\xi}_{m,l} \xi_{m,k} \chi_{m,k} = \xi_{m,k} \chi_{m,k} \mathbf{1}_{\{l=l_k\}}. \quad (4.27)$$

In heavier notation, without our simplifying conventions outlined above, we could write  $\tilde{\theta}_m$  as

$$\tilde{\theta}_m(t, x) = T_{m-1}(t, x) + \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k}(t) \tilde{\chi}_{m,k}(t, x) \nabla (T_{m-1}(t, X_{m-1,l_k}(t, \cdot))) (X_{m-1,l_k}^{-1}(t, x)) + \tilde{H}_m(t, x).$$

In order that the tensor contractions are clear, we also mention that the second term of (4.24) can be written in coordinates as

$$\sum_{k \in 2\mathbb{Z}+1} \sum_{i=1}^2 (\tilde{\chi}_{m,k})_i \xi_{m,k} \partial_{x_i} (T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1}.$$

Since, as explained above, we should consider  $T_{m-1} + \tilde{H}_m \approx \tilde{T}_{m-1}$ , which solves (4.15), the definition of our ansatz (4.24) can be compared to

$$\tilde{T}_{m-1} + \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} \tilde{\chi}_{m,k} (\nabla (\tilde{T}_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1}).$$

The latter is similar to a two-scale expansion in classical homogenization, with the role of the macroscopic function being played by  $\tilde{T}_{m-1}$  and the correctors by  $\chi_{m,k}$ . The compositions with the flows implicitly mean that the expansion is with respect to Lagrangian variables. In other words, on each time interval of size  $\tau_m''$ , we have composed with the appropriate flow  $X_{m-1,l_k}$ , written the two-scale expansion in these variables, and then composed with the inverse flow. Implicit is the assumption that, since the inverse flows  $X_{m-1,l}^{-1}$  are “slow,” the correctors should be close to the correctors for the stationary shear flows composed with the inverse flows.

We will test the validity of our ansatz (4.24) by plugging it into the left side of the equation for  $\theta_m$ , and estimating the error. This is the focus of Section 5. To prepare for this analysis, we need to obtain good regularity estimates on the “macroscopic” ingredients in the expansion: in particular, the function  $T_{m-1}$ . Indeed, as we know from classical homogenization theory, the homogenization error depends on the regularity of the macroscopic solution.

### 4.3 Regularity estimates for $T_{m-1}$

Our first order of business is to show that the equation (4.13) can be considered as a small perturbation of (4.3) with  $m$  replaced by  $m-1$ . This intuition is formalized by noting that according to (4.10)–(4.12) we have

$$T_{m-1} - \theta_{m-1} = T_{m-1}^{(N_*)} - T_{m-1}^{(0)} = \sum_{i=1}^{N_*} \underbrace{T_{m-1}^{(i)} - T_{m-1}^{(i-1)}}_{=: V_{m-1}^{(i)}}, \quad (4.28)$$

where the equation for the increment  $V_{m-1}^{(i)}$  can be found by subtracting (4.11) with  $i$  and  $i-1$ . We obtain, for each  $1 \leq i \leq N_*$ ,

$$\begin{cases} \partial_t V_{m-1}^{(i)} - \kappa_{m-1} \Delta V_{m-1}^{(i)} + \mathbf{b}_{m-1} \cdot \nabla V_{m-1}^{(i)} = \nabla \cdot (\mathbf{K}_m - \kappa_{m-1} \mathbf{I}_2 + \mathbf{s}_{m-1}) \nabla V_{m-1}^{(i-1)} & \text{in } (0, \infty) \times \mathbb{R}^2, \\ V_{m-1}^{(i)} = 0 & \text{on } \{0\} \times \mathbb{R}^2, \end{cases} \quad (4.29)$$

where for convenience we denote  $V_{m-1}^{(-1)} := 0$ , so that  $V_{m-1}^{(0)} = \theta_{m-1}$ .

In light of (4.28)–(4.29) it is apparent that we must obtain good estimates for  $\nabla \theta_{m-1}$  and  $\nabla V_{m-1}^{(i)}$ , which we achieve in Lemmas 4.1 and 4.2 below.

**Lemma 4.1** (Estimates on  $\theta_{m-1}$ ). *There exists  $1 \leq C < \infty$  such that, for every  $n \in \mathbb{N}_0$ ,*

$$\|\nabla^n \theta_{m-1}\|_{L^\infty([0,1]; L^2(\mathbb{T}^2))} + \kappa_{m-1}^{1/2} \|\nabla^{n+1} \theta_{m-1}\|_{L^2((0,1) \times \mathbb{T}^2)} \leq \|\theta_0\|_{L^2(\mathbb{T}^2)} n! (C \varepsilon_{m-1}^{-1-\gamma/2} \vee 2R_{\theta_0}^{-1})^n. \quad (4.30)$$

*Proof.* We fix a multi-index  $\alpha$  of order  $|\alpha| = n \geq 0$  and apply  $\partial^\alpha$  to both sides of (4.4) (with  $m$  replaced by  $m-1$ ) to obtain:

$$\partial_t \partial^\alpha \theta_{m-1} - \nabla \cdot (\kappa_{m-1} \mathbf{I}_2 + \phi_{m-1} \sigma) \nabla \partial^\alpha \theta_{m-1} = \nabla \cdot \sum_{\beta < \alpha} \binom{\alpha}{\beta} \partial^{\alpha-\beta} \phi_{m-1} \sigma \nabla \partial^\beta \theta_{m-1} \quad (4.31)$$

in  $(0, \infty) \times \mathbb{T}^2$ . Testing (4.31) with  $\partial^\alpha \theta_{m-1}$  and using that  $\sigma$  is skew-symmetric, we get

$$\begin{aligned} & \sup_{t \in [0,1]} \|\partial^\alpha \theta_{m-1}(t, \cdot)\|_{L^2(\mathbb{T}^2)}^2 - \|\partial^\alpha \theta_0\|_{L^2(\mathbb{T}^2)}^2 + 2\kappa_{m-1} \int_0^1 \int_{\mathbb{T}^2} |\nabla \partial^\alpha \theta_{m-1}|^2 \\ & \leq 2 \sum_{\beta < \alpha} \binom{\alpha}{\beta} \left| \int_0^1 \int_{\mathbb{T}^2} \nabla \partial^\alpha \theta_{m-1} \cdot \partial^{\alpha-\beta} \phi_{m-1} \sigma \nabla \partial^\beta \theta_{m-1} \right|. \end{aligned} \quad (4.32)$$

When  $n = 0$ , the right side of (4.31), and hence (4.32), vanishes identically. As such, we are only left to consider the case  $n \geq 1$ .

To upper bound the right side of (4.32), we split the sum into two parts: the terms involving  $\phi_{m-1}$  with  $|\alpha - \beta| \geq 2$  and the terms involving  $\phi_{m-1}$  with  $|\alpha - \beta| = 1$ . For the first group of terms, we use (2.65) and obtain

$$\begin{aligned} & \sum_{\beta < \alpha, |\alpha - \beta| \geq 2} \binom{\alpha}{\beta} \left| \int_0^1 \int_{\mathbb{T}^2} \nabla \partial^\alpha \theta_{m-1} \cdot \partial^{\alpha-\beta} \phi_{m-1} \sigma \nabla \partial^\beta \theta_{m-1} \right| \\ & \leq \kappa_{m-1} \left\| \nabla \partial^\alpha \theta_{m-1} \right\|_{L^2((0,1) \times \mathbb{T}^2)}^2 \\ & \quad + \frac{C \varepsilon_{m-1}^{2\beta}}{\kappa_{m-1}} \sum_{\beta < \alpha, |\alpha - \beta| \geq 2} \frac{|\alpha|^2}{|\beta|^2} (C \varepsilon_{m-1}^{-1})^{2|\alpha - \beta|} \left\| \nabla \partial^\beta \theta_{m-1} \right\|_{L^2((0,1) \times \mathbb{T}^2)}^2. \end{aligned} \quad (4.33)$$

For the second group, in which  $|\alpha - \beta| = 1$ , we perform an integration by parts and use the skew-symmetry of  $\sigma$  and (2.65) with  $n = 2$  to see that

$$\begin{aligned} \left| \int_0^1 \int_{\mathbb{T}^2} \nabla \partial^\alpha \theta_{m-1} \cdot \partial^{\alpha-\beta} \phi_{m-1} \sigma \nabla \partial^\beta \theta_{m-1} \right| &= \left| \int_0^1 \int_{\mathbb{T}^2} \partial^\alpha \theta_{m-1} \cdot \nabla (\partial^{\alpha-\beta} \phi_{m-1}) \cdot \sigma \nabla \partial^\beta \theta_{m-1} \right| \\ &\leq C \varepsilon_{m-1}^{\beta-2} \left\| \partial^\alpha \theta_{m-1} \right\|_{L^2((0,1) \times \mathbb{T}^2)} \left\| \nabla \partial^\beta \theta_{m-1} \right\|_{L^2((0,1) \times \mathbb{T}^2)}. \end{aligned}$$

Therefore, if  $|\alpha - \beta| = 1$ , then

$$\begin{aligned} & \sum_{\beta < \alpha, |\alpha - \beta| = 1} \binom{\alpha}{\beta} \left| \int_0^1 \int_{\mathbb{T}^2} \nabla \partial^\alpha \theta_{m-1} \cdot \partial^{\alpha-\beta} \phi_{m-1} \sigma \nabla \partial^\beta \theta_{m-1} \right| \\ & \leq C |\alpha| \varepsilon_{m-1}^{\beta-2} \left\| \partial^\alpha \theta_{m-1} \right\|_{L^2((0,1) \times \mathbb{T}^2)} \max_{\beta < \alpha, |\alpha - \beta| = 1} \left\| \nabla \partial^\beta \theta_{m-1} \right\|_{L^2((0,1) \times \mathbb{T}^2)}. \end{aligned} \quad (4.34)$$

We next insert the estimates (4.33) and (4.34) into the right side of (4.32), and appeal to (4.2) to bound the initial data term. After dividing by  $|\alpha|!$ , we obtain for  $n = |\alpha| \geq 1$  that

$$\begin{aligned} & \frac{\kappa_{m-1} \left\| \nabla \partial^\alpha \theta_{m-1} \right\|_{L^2((0,1) \times \mathbb{T}^2)}^2}{|\alpha|!^2} + \frac{\sup_{t \in [0,1]} \left\| \partial^\alpha \theta_{m-1}(t, \cdot) \right\|_{L^2(\mathbb{T}^2)}^2}{|\alpha|!^2} - \frac{\|\theta_0\|_{L^2(\mathbb{T}^2)}^2}{R_{\theta_0}^{2n}} \\ & \leq \frac{C \varepsilon_{m-1}^{2\beta}}{\kappa_{m-1}^2} \sum_{\beta < \alpha, |\alpha - \beta| \geq 2} (C \varepsilon_{m-1}^{-1})^{2|\alpha - \beta|} \frac{\kappa_{m-1} \left\| \nabla \partial^\beta \theta_{m-1} \right\|_{L^2((0,1) \times \mathbb{T}^2)}^2}{|\beta|!^2} \\ & \quad + \frac{C \varepsilon_{m-1}^{\beta-2}}{|\alpha| \kappa_{m-1}} \frac{\kappa_{m-1}^{1/2} \left\| \partial^\alpha \theta_{m-1} \right\|_{L^2((0,1) \times \mathbb{T}^2)}}{(|\alpha| - 1)!} \max_{\beta < \alpha, |\alpha - \beta| = 1} \frac{\kappa_{m-1}^{1/2} \left\| \nabla \partial^\beta \theta_{m-1} \right\|_{L^2((0,1) \times \mathbb{T}^2)}}{|\beta|!}. \end{aligned} \quad (4.35)$$

Fixing a constant  $A \geq 1$ , to be selected below (just above (4.37)), and defining

$$D_n := \frac{1}{\|\theta_0\|_{L^2(\mathbb{T}^2)}} \left( \frac{\varepsilon_{m-1}}{A} \right)^n \max_{|\alpha|=n} \left( \frac{\kappa_{m-1}^{1/2} \left\| \nabla \partial^\alpha \theta_{m-1} \right\|_{L^2((0,1) \times \mathbb{T}^2)}}{|\alpha|!} + \frac{\sup_{t \in [0,1]} \left\| \partial^\alpha \theta_{m-1}(t, \cdot) \right\|_{L^2(\mathbb{T}^2)}}{|\alpha|!} \right), \quad (4.36)$$

we then take the maximum of (4.35) over all multi-indices  $\alpha$  with  $|\alpha| = n$  and rearranging the resulting expression, using also elementary bounds for multinomial coefficients and factorials, to obtain

$$D_n^2 - \left( \frac{\varepsilon_{m-1}}{A R_{\theta_0}} \right)^{2n} \leq \frac{C \varepsilon_{m-1}^{2\beta}}{\kappa_{m-1}^2} \sum_{k=0}^{n-2} \left( \frac{C}{A} \right)^{2(n-k)} D_k^2 + \frac{C \varepsilon_{m-1}^\beta}{n \kappa_{m-1} A^2} D_{n-1}^2.$$

Using the upper and lower bounds for  $\kappa_{m-1}$  from (3.45), for  $n \geq 1$  we obtain from the above estimate that

$$D_n^2 \leq \left( \frac{\varepsilon_{m-1}}{AR_{\theta_0}} \right)^{2n} + \frac{C}{\varepsilon_{m-1}^{2\gamma}} \sum_{k=0}^{n-2} \left( \frac{C}{A} \right)^{2(n-k)} D_k^2 + \frac{C}{\varepsilon_{m-1}^\gamma A^2} D_{n-1}^2.$$

Note that when  $n = 0$  only the first term on the right side of the above estimate is present. If we choose  $A$  by

$$A := \max \left\{ 1, 4C^{3/2}, (4C)^{1/2}, 2\varepsilon_{m-1}^{1+\gamma/2} R_{\theta_0}^{-1} \right\},$$

then we obtain

$$D_n \leq 2^{-n} (\varepsilon_{m-1}^{-\gamma/2})^n + 2^{-2} \mathbf{1}_{n \geq 2} \varepsilon_{m-1}^{-\gamma} \max\{D_0, \dots, D_{n-2}\} + 2^{-2} \mathbf{1}_{n \geq 1} \varepsilon_{m-1}^{-\gamma/2} D_{n-1}. \quad (4.37)$$

Iterating this inequality, we discover that, for every  $n \in \mathbb{N}_0$ ,

$$D_n \leq (\varepsilon_{m-1}^{-\gamma/2})^n.$$

Recalling (4.36) and the notation (1.22), the above estimate implies

$$\sup_{t \in [0,1]} \|\nabla^n \theta_{m-1}(t, \cdot)\|_{L^2(\mathbb{T}^2)} + \kappa_{m-1}^{1/2} \|\nabla^{n+1} \theta_{m-1}\|_{L^2((0,1) \times \mathbb{T}^2)} \leq \|\theta_0\|_{L^2(\mathbb{T}^2)} n! (A \varepsilon_{m-1}^{-1-\gamma/2})^n.$$

In view of our choice of  $A$ , the above estimate gives (4.30).  $\square$

Next, we aim to obtain similar regularity estimates for  $\{V_{m-1}^{(i)}\}_{i=1}^{N_*}$ . Since the equation (4.29) contains the matrix  $\mathbf{s}_{m-1}$ , we first need to obtain suitable estimates for this function. We show  $\mathbf{s}_{m-1}$  is small relative to  $\kappa_{m-1}$ , and the scale of its spatial oscillations are large compared to  $\varepsilon_m$ . Indeed, by (2.67) and (2.14), we have that the term defined in (4.7) satisfies

$$\begin{aligned} \|\mathbf{s}_{m-1}\|_{L^\infty(\mathbb{R} \times \mathbb{T}^2)} &\leq \kappa_{m-1} \sup_{l \in \mathbb{Z}} \|\hat{\xi}_{m,l}(\nabla X_{m-1,l} - \mathbf{I}_2)\|_{L^\infty(\mathbb{R} \times \mathbb{T}^2)} \leq C \kappa_{m-1} \varepsilon_{m-1}^{2\delta} \\ &= C a_{m-1} \varepsilon_{m-1}^{2+\gamma+2\delta} = C \varepsilon_{m-1}^{\beta+\gamma+2\delta}. \end{aligned} \quad (4.38)$$

More generally, we have the following bound on the higher-order spatial derivatives of  $\mathbf{s}_{m-1}$ :

$$\llbracket \mathbf{s}_{m-1} \rrbracket_{C\varepsilon_{m-1}^{-1}} \leq C \kappa_{m-1} = C \varepsilon_{m-1}^{\beta+\gamma}. \quad (4.39)$$

The estimate (4.39) is a consequence of (2.68) and Proposition B.6.

Comparing (4.39) to (2.65) and (2.66), we see that  $\mathbf{s}_{m-1}$  is smaller than  $\phi_m$  by a factor of  $\varepsilon^{\gamma+2\delta}$  while having the same analyticity radius, and is smaller in size than  $\kappa_{m-1} \simeq a_m \varepsilon_m^{2+\gamma}$  by a factor of  $\varepsilon_{m-1}^{2\delta}$ . Our next goal is to use this fact to show that the bounds obeyed by  $V_{m-1}^{(i)}$  are better than those satisfied by  $\theta_{m-1}$ , by a factor of at least  $\varepsilon_{m-1}^{2\delta}$ .

**Lemma 4.2** (Estimates on  $V_{m-1}^{(i)}$ ). *There exists a constant  $C_0 < \infty$  such that, if  $\varepsilon_{m-1}$  is small enough that*

$$C_0^3 \varepsilon_{m-1}^{2\delta} (1 \vee \varepsilon_{m-1}^{2+\gamma} R_{\theta_0}^{-2}) \leq 1, \quad (4.40)$$

*then, for every  $n, i \in \mathbb{N}_0$ , we have the estimate*

$$\begin{aligned} \frac{1}{(n+2i)!} (C_0 \varepsilon_{m-1}^{-1-\gamma/2} \vee C_0 R_{\theta_0}^{-1})^{-n} &\left( \|\nabla^n V_{m-1}^{(i)}\|_{L^\infty([0,1]; L^2(\mathbb{T}^2))} + \kappa_{m-1}^{1/2} \|\nabla^{n+1} V_{m-1}^{(i)}\|_{L^2((0,1) \times \mathbb{T}^2)} \right) \\ &\leq \mathsf{A}_{m-1,i} \|\theta_0\|_{L^2(\mathbb{T}^2)}, \end{aligned} \quad (4.41)$$

where we have defined the amplitude coefficients appearing in (4.41) by

$$\mathsf{A}_{m-1,i} := \mathbf{1}_{\{i=0\}} + \mathbf{1}_{\{i \in \{1,2\}\}} \left( C_0^3 \varepsilon_{m-1}^{2\delta} (1 \vee \varepsilon_{m-1}^{2+\gamma} R_{\theta_0}^{-2}) \right) + \mathbf{1}_{\{i \geq 3\}} \left( C_0^3 \varepsilon_{m-1}^{2\delta} (1 \vee \varepsilon_{m-1}^{2+\gamma} R_{\theta_0}^{-2}) \right)^{\frac{i}{2}}. \quad (4.42)$$

*Proof.* The proof of Lemma 4.2 is recursive in  $i \geq 0$ , and closely follows the proof of Lemma 4.1. Since  $V_{m-1}^{(0)} = \theta_{m-1}$ , the bound (4.1) establishes the inductive step, namely (4.41) for  $i = 0$ , as long as we ensure  $C_0 \geq C_{(4.30)}$ .

Next, assume that (4.41) with  $i$  replaced by  $i - 1$ . Comparing the  $V_{m-1}^{(i)}$  evolution (4.29) and the  $\theta_{m-1}$  evolution in (4.4) (with  $m$  replaced by  $m - 1$ ), we see that the only difference is due to the forcing term  $\nabla \cdot (\mathbf{K}_m - \kappa_{m-1} \mathbf{I}_2 + \mathbf{s}_{m-1}) \nabla V_{m-1}^{(i-1)}$ , and the fact that the initial data for  $V_{m-1}^{(i)}$  vanishes identically. As such, since  $\mathbf{K}_m$  is only a function of time, (4.32) becomes

$$\begin{aligned} & \sup_{t \in [0,1]} \|\partial^\alpha V_{m-1}^{(i)}(t, \cdot)\|_{L^2(\mathbb{T}^2)}^2 + 2\kappa_{m-1} \int_0^1 \int_{\mathbb{T}^2} |\nabla \partial^\alpha V_{m-1}^{(i)}|^2 \\ & \leq 2 \sum_{\beta < \alpha} \binom{\alpha}{\beta} \left| \int_0^1 \int_{\mathbb{T}^2} \nabla \partial^\alpha V_{m-1}^{(i)} \cdot \partial^{\alpha-\beta} \phi_{m-1} \sigma \nabla \partial^\beta V_{m-1}^{(i)} \right| \\ & \quad + 2 \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \left| \int_0^1 \int_{\mathbb{T}^2} \nabla \partial^\alpha V_{m-1}^{(i)} \cdot \partial^{\alpha-\beta} \mathbf{s}_{m-1} \nabla \partial^\beta V_{m-1}^{(i-1)} \right| \\ & \quad + 2 \left| \int_0^1 \int_{\mathbb{T}^2} \nabla \partial^\alpha V_{m-1}^{(i)} \cdot (\mathbf{K}_m - \kappa_{m-1} \mathbf{I}_2) \nabla \partial^\alpha V_{m-1}^{(i-1)} \right| \\ & =: \mathsf{Err}_1 + \mathsf{Err}_2 + \mathsf{Err}_3. \end{aligned} \quad (4.43)$$

The first term on the right side of (4.43),  $\mathsf{Err}_1$ , is estimated in exactly the same fashion as (4.33) and (4.34), resulting in the estimate

$$\begin{aligned} \mathsf{Err}_1 & \leq \frac{\kappa_{m-1}}{4} \|\nabla \partial^\alpha V_{m-1}^{(i)}\|_{L^2((0,1) \times \mathbb{T}^2)}^2 \\ & \quad + \frac{C \varepsilon_{m-1}^{2\beta}}{\kappa_{m-1}} \sum_{\beta < \alpha, |\alpha-\beta| \geq 2} \frac{|\alpha|!^2}{|\beta|!^2} (C \varepsilon_{m-1}^{-1})^{2|\alpha-\beta|} \|\nabla \partial^\beta V_{m-1}^{(i)}\|_{L^2((0,1) \times \mathbb{T}^2)}^2 \\ & \quad + C |\alpha| \varepsilon_{m-1}^{\beta-2} \|\partial^\alpha V_{m-1}^{(i)}\|_{L^2((0,1) \times \mathbb{T}^2)} \max_{\beta < \alpha, |\alpha-\beta|=1} \|\nabla \partial^\beta V_{m-1}^{(i)}\|_{L^2((0,1) \times \mathbb{T}^2)}. \end{aligned} \quad (4.44)$$

In order to estimate the second term on the right side of (4.43),  $\mathsf{Err}_2$ , we appeal to the  $\mathbf{s}_{m-1}$  bounds (4.38)–(4.39) and to the inductive estimate for  $V_{m-1}^{(i-1)}$  provided by (4.41). We arrive at

$$\begin{aligned} \mathsf{Err}_2 & \leq \frac{\kappa_{m-1}}{4} \|\nabla \partial^\alpha V_{m-1}^{(i)}\|_{L^2((0,1) \times \mathbb{T}^2)}^2 \\ & \quad + C \mathsf{A}_{m-1,i-1}^2 \|\theta_0\|_{L^2(\mathbb{T}^2)}^2 \sum_{\beta < \alpha} \frac{|\alpha|!^2 (|\beta| + 2i - 2)!^2}{|\beta|!^2} (C \varepsilon_{m-1}^{-1})^{2|\alpha-\beta|} (C_0 \varepsilon_{m-1}^{-1-\gamma/2} \vee C_0 R_{\theta_0}^{-1})^{2|\beta|} \\ & \quad + C \varepsilon_{m-1}^{4\delta} \mathsf{A}_{m-1,i-1}^2 \|\theta_0\|_{L^2(\mathbb{T}^2)}^2 (|\alpha| + 2i - 2)!^2 (C_0 \varepsilon_{m-1}^{-1-\gamma/2} \vee C_0 R_{\theta_0}^{-1})^{2|\alpha|}. \end{aligned} \quad (4.45)$$

Bounding the last term on the right side of (4.43),  $\mathsf{Err}_3$ , requires more care. First, recalling (3.42) we note that

$$\mathbf{K}_m - \kappa_{m-1} \mathbf{I}_2 = \mathbf{K}_m - \bar{\mathbf{K}}_m = (\mathbf{K}_m - \langle\langle \mathbf{K}_m \rangle\rangle) + (\langle\langle \mathbf{K}_m \rangle\rangle - \bar{\mathbf{K}}_m). \quad (4.46)$$

By appealing to (3.29), (3.39), and (3.46), the term  $|\langle\langle \mathbf{K}_m \rangle\rangle - \bar{\mathbf{K}}_m|$  may be made arbitrarily small, which may be combined with (4.41) at level  $i-1$  to deduce the bound

$$\begin{aligned}
\mathsf{Err}_3 &\leq 2 \left| \int_0^1 \int_{\mathbb{T}^2} \nabla \partial^\alpha V_{m-1}^{(i)} \cdot (\mathbf{K}_m - \langle\langle \mathbf{K}_m \rangle\rangle) \nabla \partial^\alpha V_{m-1}^{(i-1)} \right| \\
&\quad + \frac{C a_m^2 \varepsilon_m^4}{\kappa_m} \left( \frac{\varepsilon_m^2}{\kappa_m \tau_m} \right)^{N_*} \|\nabla \partial^\alpha V_{m-1}^{(i)}\|_{L^2((0,1) \times \mathbb{T}^2)} \|\nabla \partial^\alpha V_{m-1}^{(i-1)}\|_{L^2((0,1) \times \mathbb{T}^2)} \\
&\leq 2 \left| \int_0^1 \int_{\mathbb{T}^2} \nabla \partial^\alpha V_{m-1}^{(i)} \cdot (\mathbf{K}_m - \langle\langle \mathbf{K}_m \rangle\rangle) \nabla \partial^\alpha V_{m-1}^{(i-1)} \right| \\
&\quad + \frac{\kappa_{m-1}}{4} \|\nabla \partial^\alpha V_{m-1}^{(i)}\|_{L^2((0,1) \times \mathbb{T}^2)}^2 \\
&\quad + C (C \varepsilon_{m-1}^{2\delta})^{2N_*} A_{m-1,i-1}^2 \|\theta_0\|_{L^2(\mathbb{T}^2)}^2 (|\alpha| + 2i - 2)!^2 (C_0 \varepsilon_{m-1}^{-1-\gamma/2} \vee C_0 R_{\theta_0}^{-1})^{2|\alpha|}. \quad (4.47)
\end{aligned}$$

It thus remains to estimate the first term on the right side of (4.47). For this purpose, we recall from (3.26) that

$$\mathbf{K}_m(t) - \langle\langle \mathbf{K}_m \rangle\rangle = \sum_{n=0}^{N_*-1} \langle\langle \mathbf{j}_{m,n} \rangle\rangle (L_{m,n}(t) - \langle\langle L_{m,n} \rangle\rangle)$$

is a zero-mean symmetric matrix, which is  $\tau_m''$ -periodic in time. As such, we may write

$$\mathbf{K}_m(t) - \langle\langle \mathbf{K}_m \rangle\rangle = \partial_t \mathbf{Q}_m(t) = \mathcal{D}_{t,m-1} \mathbf{Q}_m(t), \quad (4.48)$$

where

$$\|\mathbf{Q}_m\|_{L^\infty([0,1])} \leq C \kappa_{m-1} \tau_m'', \quad \text{and} \quad \mathbf{Q}_m(0) = \mathbf{Q}_m(1) = 0. \quad (4.49)$$

Using the above two displays, we may integrate by parts (since  $\nabla \cdot \mathbf{b}_{m-1} = 0$ , the  $L^2$ -adjoint of the operator  $\mathbf{b}_{m-1} \cdot \nabla$  is the operator  $-\mathbf{b}_{m-1} \cdot \nabla$ ) and deduce that

$$\begin{aligned}
&\int_0^1 \int_{\mathbb{T}^2} \nabla \partial^\alpha V_{m-1}^{(i)} \cdot (\mathbf{K}_m - \langle\langle \mathbf{K}_m \rangle\rangle) \nabla \partial^\alpha V_{m-1}^{(i-1)} \\
&= - \int_0^1 \int_{\mathbb{T}^2} \nabla \partial^\alpha V_{m-1}^{(i)} \cdot \mathbf{Q}_m \mathcal{D}_{t,m-1} \nabla \partial^\alpha V_{m-1}^{(i-1)} - \int_0^1 \int_{\mathbb{T}^2} \mathcal{D}_{t,m-1} \nabla \partial^\alpha V_{m-1}^{(i)} \cdot \mathbf{Q}_m \nabla \partial^\alpha V_{m-1}^{(i-1)} \\
&= \int_0^1 \int_{\mathbb{T}^2} \nabla \partial^\alpha V_{m-1}^{(i)} \cdot \mathbf{Q}_m \nabla \mathbf{b}_{m-1} \cdot \nabla \partial^\alpha V_{m-1}^{(i-1)} + \int_0^1 \int_{\mathbb{T}^2} \nabla \mathbf{b}_{m-1} \cdot \nabla \partial^\alpha V_{m-1}^{(i)} \cdot \mathbf{Q}_m \nabla \partial^\alpha V_{m-1}^{(i-1)} \\
&\quad - \int_0^1 \int_{\mathbb{T}^2} \nabla \partial^\alpha V_{m-1}^{(i)} \cdot \mathbf{Q}_m \nabla \mathcal{D}_{t,m-1} \partial^\alpha V_{m-1}^{(i-1)} - \int_0^1 \int_{\mathbb{T}^2} \nabla \mathcal{D}_{t,m-1} \partial^\alpha V_{m-1}^{(i)} \cdot \mathbf{Q}_m \nabla \partial^\alpha V_{m-1}^{(i-1)} \\
&=: \mathsf{Err}_{4,1} + \mathsf{Err}_{4,2} + \mathsf{Err}_{4,3} + \mathsf{Err}_{4,4}. \quad (4.50)
\end{aligned}$$

Using the  $\nabla \mathbf{b}_{m-1}$  estimate in (2.72), the  $\mathbf{Q}_m$  bound in (4.49), the estimate  $\tau_m'' \varepsilon_{m-1}^{\beta-2} \leq C \varepsilon_{m-1}^{2\delta}$ , we obtain  $|\nabla \mathbf{b}_{m-1}| |\mathbf{Q}_m| \leq C \kappa_{m-1} \varepsilon_{m-1}^{2\delta}$ , and so, by also using the inductive estimate for  $\nabla \partial^\alpha V_{m-1}^{(i-1)}$  in (4.2), we obtain

$$\begin{aligned}
|\mathsf{Err}_{4,1}| + |\mathsf{Err}_{4,2}| &\leq \frac{\kappa_{m-1}}{24} \|\nabla \partial^\alpha V_{m-1}^{(i)}\|_{L^2((0,1) \times \mathbb{T}^2)}^2 \\
&\quad + C \varepsilon_{m-1}^{4\delta} A_{m-1,i-1}^2 \|\theta_0\|_{L^2(\mathbb{T}^2)}^2 (|\alpha| + 2i - 2)!^2 (C_0 \varepsilon_{m-1}^{-1-\gamma/2} \vee C_0 R_{\theta_0}^{-1})^{2|\alpha|}. \quad (4.51)
\end{aligned}$$

Next, we bound the  $\text{Err}_{4,3}$  and  $\text{Err}_{4,4}$  terms appearing on the right side of (4.50). For this purpose, we note that (4.29) gives

$$\begin{aligned} \nabla \mathcal{D}_{t,m-1} \partial^\alpha V_{m-1}^{(i)} &= \kappa_{m-1} \Delta \nabla \partial^\alpha V_{m-1}^{(i)} - \nabla \sum_{\beta < \alpha} \binom{\alpha}{\beta} \partial^{\alpha-\beta} \mathbf{b}_{m-1} \cdot \nabla \partial^\beta V_{m-1}^{(i)} \\ &\quad + \mathbf{1}_{i \geq 1} \nabla \nabla \cdot (\mathbf{K}_m - \kappa_{m-1} \mathbf{I}_2) \nabla \partial^\alpha V_{m-1}^{(i-1)} \\ &\quad + \mathbf{1}_{i \geq 1} \nabla \nabla \cdot \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \partial^{\alpha-\beta} \mathbf{s}_{m-1} \nabla \partial^\beta V_{m-1}^{(i-1)}. \end{aligned} \quad (4.52)$$

Using (4.52) we first bound the more difficult term,  $\text{Err}_4$ . The additional complication arises from the fact that the first line on the right side of (4.52) contains terms with 2, and respectively 1, additional derivatives on top of  $\nabla \partial^\alpha$ , and this apparently prevents us from closing our estimates; this is however not an issue, as these derivatives may be integrated by parts onto the  $T_{m-1}^{(i-1)}$  term, for which we have already estimated all the space derivatives (including those of order  $n+3$ ). To be precise, (4.52) allows us to rewrite

$$\begin{aligned} \text{Err}_{4,4} &= -\kappa_{m-1} \int_0^1 \int_{\mathbb{T}^2} \nabla \partial^\alpha V_{m-1}^{(i)} \cdot \mathbf{Q}_m \Delta \nabla \partial^\alpha V_{m-1}^{(i-1)} \\ &\quad - \sum_{\beta < \alpha} \binom{\alpha}{\beta} \int_0^1 \int_{\mathbb{T}^2} \partial^{\alpha-\beta} \mathbf{b}_{m-1} \cdot \nabla \partial^\beta V_{m-1}^{(i)} \cdot \mathbf{Q}_m \nabla \nabla \partial^\alpha V_{m-1}^{(i-1)} \\ &\quad + \mathbf{1}_{i \geq 1} \int_0^1 \int_{\mathbb{T}^2} \left( \nabla \cdot (\mathbf{K}_m - \kappa_{m-1} \mathbf{I}_2) \nabla \partial^\alpha V_{m-1}^{(i-1)} \right) \left( \nabla \cdot \mathbf{Q}_m \nabla \partial^\alpha V_{m-1}^{(i-1)} \right) \\ &\quad - \mathbf{1}_{i \geq 1} \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \int_0^1 \int_{\mathbb{T}^2} \nabla \nabla \cdot (\partial^{\alpha-\beta} \mathbf{s}_{m-1} \nabla \partial^\beta V_{m-1}^{(i-1)}) \cdot \mathbf{Q}_m \nabla \partial^\alpha V_{m-1}^{(i-1)}. \end{aligned} \quad (4.53)$$

Note that the ‘‘gain’’ we expect for  $i=1$  (see (4.42)) is larger than for  $i \geq 2$ , and because of that, special care must be devoted to the third term on the right side of (4.53), when  $i=1$ . Using (4.46) and (4.48), and recalling that  $V_{m-1}^{(0)} = \theta_{m-1}$ , we rewrite

$$\begin{aligned} \mathbf{1}_{i=1} \int_0^1 \int_{\mathbb{T}^2} &\left( \nabla \cdot (\mathbf{K}_m - \kappa_{m-1} \mathbf{I}_2) \nabla \partial^\alpha V_{m-1}^{(i-1)} \right) \left( \nabla \cdot \mathbf{Q}_m \nabla \partial^\alpha V_{m-1}^{(i-1)} \right) \\ &= \int_0^1 \int_{\mathbb{T}^2} \left( \nabla \cdot (\langle\langle \mathbf{K}_m \rangle\rangle - \bar{\mathbf{K}}_m) \nabla \partial^\alpha \theta_{m-1} \right) \left( \nabla \cdot \mathbf{Q}_m \nabla \partial^\alpha \theta_{m-1} \right) \\ &\quad + \int_0^1 \int_{\mathbb{T}^2} \left( \mathcal{D}_{t,m-1} \mathbf{Q}_m : \nabla^2 \partial^\alpha \theta_{m-1} \right) \left( \mathbf{Q}_m : \nabla^2 \partial^\alpha \theta_{m-1} \right) \\ &= \int_0^1 \int_{\mathbb{T}^2} \left( \nabla \cdot (\langle\langle \mathbf{K}_m \rangle\rangle - \bar{\mathbf{K}}_m) \nabla \partial^\alpha \theta_{m-1} \right) \left( \nabla \cdot \mathbf{Q}_m \nabla \partial^\alpha \theta_{m-1} \right) \\ &\quad - \int_0^1 \int_{\mathbb{T}^2} \left( \mathbf{Q}_m : \mathcal{D}_{t,m-1} \nabla^2 \partial^\alpha \theta_{m-1} \right) \left( \mathbf{Q}_m : \nabla^2 \partial^\alpha \theta_{m-1} \right). \end{aligned} \quad (4.54)$$

The last term in the above expression may then be rewritten upon noting that  $\mathcal{D}_{t,m-1} \theta_{m-1} = \kappa_{m-1} \Delta \theta_{m-1}$ , and therefore

$$\begin{aligned} \mathcal{D}_{t,m-1} \partial^{\alpha+e_r+e_p} \theta_{m-1} &= \kappa_{m-1} \Delta \partial^{\alpha+e_r+e_p} \theta_{m-1} + \sum_{\beta < \alpha+e_r+e_p} \binom{\alpha+e_r+e_p}{\beta} \partial^{\alpha+e_r+e_p-\beta} \mathbf{b}_{m-1} \cdot \nabla \partial^\beta \theta_{m-1}. \end{aligned}$$

Combining the above identity with (4.54) and (4.53), and appealing to the  $\mathbf{b}_{m-1}$  estimate in (2.72), the  $\mathbf{Q}_m$  bound in (4.49), the  $\mathbf{s}_{m-1}$  estimate in (4.39), the inductive bound (4.2) at level  $i-1$ , and to (3.29) and (3.46) for  $i=1$ , we deduce

$$\begin{aligned}
|\text{Err}_{4,4}| &\leq \frac{\kappa_{m-1}}{24} \|\nabla \partial^\alpha V_{m-1}^{(i)}\|_{L^2((0,1) \times \mathbb{T}^2)}^2 \\
&\quad + C(C\kappa_{m-1}\tau_m'')^2 \mathbf{A}_{m-1,i-1}^2 \|\theta_0\|_{L^2(\mathbb{T}^2)}^2 (|\alpha|+2i)!^2 (C_0 \varepsilon_{m-1}^{-1-\gamma/2} \vee C_0 R_{\theta_0}^{-1})^{2|\alpha|+4} \\
&\quad + C \sum_{\beta < \alpha} \frac{|\alpha|!(|\alpha|+2i-1)!}{|\beta|!|\alpha-\beta|} (C \varepsilon_{m-1}^{-1})^{|\alpha-\beta|} \kappa_{m-1}^{1/2} \|\nabla \partial^\beta V_{m-1}^{(i)}\|_{L^2((0,1) \times \mathbb{T}^2)} \\
&\quad \times \varepsilon_{m-1}^{\beta-1} \tau_m'' \|\theta_0\|_{L^2(\mathbb{T}^2)} \mathbf{A}_{m-1,i-1} (C_0 \varepsilon_{m-1}^{-1-\gamma/2} \vee C_0 R_{\theta_0}^{-1})^{|\alpha|+1} \\
&\quad + C \mathbf{1}_{i \geq 2} \kappa_{m-1} \tau_m'' \mathbf{A}_{m-1,i-1}^2 \|\theta_0\|_{L^2(\mathbb{T}^2)}^2 (|\alpha|+2i-1)!^2 (C_0 \varepsilon_{m-1}^{-1-\gamma/2} \vee C_0 R_{\theta_0}^{-1})^{2|\alpha|+2} \\
&\quad + \mathbf{1}_{i=1} C \kappa_{m-1} \tau_m'' (C \varepsilon_{m-1}^{2\delta})^{N_*} \|\theta_0\|_{L^2(\mathbb{T}^2)}^2 (|\alpha|+1)!^2 (C_0 \varepsilon_{m-1}^{-1-\gamma/2} \vee C_0 R_{\theta_0}^{-1})^{2|\alpha|+2} \\
&\quad + \mathbf{1}_{i=1} (C \kappa_{m-1} \tau_m'')^2 \|\theta_0\|_{L^2(\mathbb{T}^2)}^2 (|\alpha|+1)! (|\alpha|+3)! (C_0 \varepsilon_{m-1}^{-1-\gamma/2} \vee C_0 R_{\theta_0}^{-1})^{2|\alpha|+4} \\
&\quad + \mathbf{1}_{i=1} (C \kappa_{m-1} \tau_m'')^2 \kappa_{m-1}^{-1} \|\theta_0\|_{L^2(\mathbb{T}^2)}^2 (|\alpha|+1)! (C_0 \varepsilon_{m-1}^{-1-\gamma/2} \vee C_0 R_{\theta_0}^{-1})^{|\alpha|+1} \\
&\quad \times \sum_{|\beta| \leq |\alpha|+1} \frac{(|\alpha|+2)!}{|\alpha-\beta|+2} (C \varepsilon_{m-1}^{\beta-2}) (C \varepsilon_{m-1}^{-1})^{|\alpha-\beta|+1} (C_0 \varepsilon_{m-1}^{-1-\gamma/2} \vee C_0 R_{\theta_0}^{-1})^{|\beta|} \\
&\quad + C \mathbf{1}_{i \geq 1} \sum_{\beta < \alpha} \frac{|\alpha|!(|\alpha|+2i)! (|\beta|+2i-2)!}{|\beta|! (|\alpha-\beta|+1)^2} (C \varepsilon_{m-1}^{-1})^{|\alpha-\beta|} \\
&\quad \times (\kappa_{m-1} \tau_m'') \|\theta_0\|_{L^2(\mathbb{T}^2)}^2 \mathbf{A}_{m-1,i-1}^2 (C_0 \varepsilon_{m-1}^{-1-\gamma/2} \vee C_0 R_{\theta_0}^{-1})^{|\alpha|+|\beta|+2} \\
&\quad + C \mathbf{1}_{i \geq 1} (|\alpha|+2i)! (|\alpha|+2i-2)! \\
&\quad \times (\kappa_{m-1} \tau_m'' \varepsilon_{m-1}^{2\delta}) \|\theta_0\|_{L^2(\mathbb{T}^2)}^2 \mathbf{A}_{m-1,i-1}^2 (C_0 \varepsilon_{m-1}^{-1-\gamma/2} \vee C_0 R_{\theta_0}^{-1})^{2|\alpha|+2}. \tag{4.55}
\end{aligned}$$

Returning to (4.50), we are left to consider the term  $\text{Err}_{4,3}$ . The difference between this term and  $\text{Err}_{4,4}$  is that  $\mathcal{D}_{t,m-1}$  acts on  $V_{m-1}^{(i-1)}$  instead of  $V_{m-1}^{(i)}$ , and as such we need to appeal to the identity (4.52) with  $i$  replaced by  $i-1$ . We do not however need to integrate by parts terms with a derivative count larger than  $|\alpha|+1$  because they occur only on  $V_{m-1}^{(i-1)}$  and  $V_{m-1}^{(i-2)}$ . As such, similarly to (4.53) we may rewrite

$$\begin{aligned}
\text{Err}_{4,3} &= -\kappa_{m-1} \int_0^1 \int_{\mathbb{T}^2} \nabla \partial^\alpha V_{m-1}^{(i)} \cdot \mathbf{Q}_m \Delta \nabla \partial^\alpha V_{m-1}^{(i-1)} \\
&\quad + \sum_{\beta < \alpha} \binom{\alpha}{\beta} \int_0^1 \int_{\mathbb{T}^2} \nabla \partial^\alpha V_{m-1}^{(i)} \cdot \mathbf{Q}_m \nabla (\partial^{\alpha-\beta} \mathbf{b}_{m-1} \cdot \nabla \partial^\beta V_{m-1}^{(i-1)}) \\
&\quad - \mathbf{1}_{i \geq 2} \int_0^1 \int_{\mathbb{T}^2} \nabla \partial^\alpha V_{m-1}^{(i)} \cdot \mathbf{Q}_m \nabla \nabla \cdot (\mathbf{K}_m - \kappa_{m-1} \mathbf{I}_2) \nabla \partial^\alpha V_{m-1}^{(i-2)} \\
&\quad - \mathbf{1}_{i \geq 2} \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \int_0^1 \int_{\mathbb{T}^2} \nabla \partial^\alpha V_{m-1}^{(i)} \cdot \mathbf{Q}_m \nabla \nabla \cdot (\partial^{\alpha-\beta} \mathbf{s}_{m-1} \nabla \partial^\beta V_{m-1}^{(i-2)}),
\end{aligned}$$

and similarly to (4.55) we may bound

$$\begin{aligned}
|\text{Err}_{4,3}| &\leq \frac{\kappa_{m-1}}{24} \|\nabla \partial^\alpha V_{m-1}^{(i)}\|_{L^2((0,1) \times \mathbb{T}^2)}^2 \\
&\quad + C(C\kappa_{m-1}\tau_m'')^2 \mathsf{A}_{m-1,i-1}^2 \|\theta_0\|_{L^2(\mathbb{T}^2)}^2 (|\alpha| + 2i)!^2 (C_0 \varepsilon_{m-1}^{-1-\gamma/2} \vee C_0 R_{\theta_0}^{-1})^{2|\alpha|+4} \\
&\quad + C \sum_{\beta < \alpha} \frac{|\alpha|!^2 (|\beta| + 2i - 1)!^2}{|\beta|!^2} (C \varepsilon_{m-1}^{-1})^{2|\alpha-\beta|} \\
&\quad \times (\varepsilon_{m-1}^{\beta-1} \tau_m'')^2 \|\theta_0\|_{L^2(\mathbb{T}^2)}^2 \mathsf{A}_{m-1,i-1}^2 (C_0 \varepsilon_{m-1}^{-1-\gamma/2} \vee C_0 R_{\theta_0}^{-1})^{2|\beta|+2} \\
&\quad + C \mathbf{1}_{i \geq 2} (\kappa_{m-1}\tau_m'')^2 \mathsf{A}_{m-1,i-2}^2 \|\theta_0\|_{L^2(\mathbb{T}^2)}^2 (|\alpha| + 2i - 2)!^2 (C_0 \varepsilon_{m-1}^{-1-\gamma/2} \vee C_0 R_{\theta_0}^{-1})^{2|\alpha|+4} \\
&\quad + C \mathbf{1}_{i \geq 2} \sum_{\beta < \alpha} \frac{|\alpha|!^2 (|\beta| + 2i - 2)!^2}{|\beta|!^2} (C \varepsilon_{m-1}^{-1})^{2|\alpha-\beta|} \\
&\quad \times (\kappa_{m-1}\tau_m'')^2 \|\theta_0\|_{L^2(\mathbb{T}^2)}^2 \mathsf{A}_{m-1,i-2}^2 (C_0 \varepsilon_{m-1}^{-1-\gamma/2} \vee C_0 R_{\theta_0}^{-1})^{2|\beta|+4} \\
&\quad + C \mathbf{1}_{i \geq 2} (|\alpha| + 2i - 2)!^2 \\
&\quad \times (\kappa_{m-1}\tau_m'' \varepsilon_{m-1}^{2\delta})^2 \|\theta_0\|_{L^2(\mathbb{T}^2)}^2 \mathsf{A}_{m-1,i-2}^2 (C_0 \varepsilon_{m-1}^{-1-\gamma/2} \vee C_0 R_{\theta_0}^{-1})^{2|\alpha|+4}. \tag{4.56}
\end{aligned}$$

Next, in analogy to (4.36) we define

$$\begin{aligned}
D_n^{(i)} &:= \frac{(C_0 \varepsilon_{m-1}^{-1-\gamma/2} \vee C_0 R_{\theta_0}^{-1})^{-n}}{\mathsf{A}_{m-1,i} \|\theta_0\|_{L^2(\mathbb{T}^2)}} \\
&\quad \times \max_{|\alpha|=n} \left( \frac{\kappa_{m-1}^{1/2} \|\nabla \partial^\alpha V_{m-1}^{(i)}\|_{L^2((0,1) \times \mathbb{T}^2)}}{(|\alpha| + 2i)!} + \frac{\sup_{t \in [0,1]} \|\partial^\alpha V_{m-1}^{(i)}(t, \cdot)\|_{L^2(\mathbb{T}^2)}}{(|\alpha| + 2i)!} \right), \tag{4.57}
\end{aligned}$$

so that proving (4.41) amounts to showing that  $D_n^{(i)} \leq 1$ . To achieve this bound, we combine (4.43), (4.44), (4.45), (4.47), (4.51), (4.56), and absorb the appropriate term on the left side of the inequality. By also using the parameter inequalities (2.16), (3.45), the fact that  $\gamma \geq 4\delta$  (a consequence of (2.5), (2.7), and  $q > 1$ ), and upon denoting<sup>13</sup>

$$\begin{aligned}
\mathcal{F} &:= \varepsilon_{m-1}^{1+\gamma/2} (C_0 \varepsilon_{m-1}^{-1-\gamma/2} \vee C_0 R_{\theta_0}^{-1}) = C_0 (1 \vee \varepsilon_{m-1}^{1+\gamma/2} R_{\theta_0}^{-1}) \geq C_0 \geq 1, \\
\mathcal{G} &:= (C \varepsilon_{m-1}^{-1}) (C_0 \varepsilon_{m-1}^{-1-\gamma/2} \vee C_0 R_{\theta_0}^{-1})^{-1} = (C C_0^{-1}) (\varepsilon_{m-1}^{\gamma/2} \wedge \varepsilon_{m-1}^{-1} R_{\theta_0}) \leq 1/2,
\end{aligned}$$

for all  $n \geq 0$  and  $i \geq 1$  we arrive at

$$\begin{aligned}
(D_n^{(i)})^2 &\leq \frac{C}{\mathcal{F}^2} (D_{n-1}^{(i)})^2 + \frac{C}{\mathcal{F}^4} \sum_{k=0}^{n-2} 2^{-2(n-2-k)} (D_k^{(i)})^2 + \frac{C \varepsilon_{m-1}^{2\delta} \mathsf{A}_{m-1,i-1}}{\mathsf{A}_{m-1,i}} \sum_{k=0}^{n-1} 2^{-(n-1-k)} D_k^{(i)} \\
&\quad + \frac{C \mathsf{A}_{m-1,i-1}^2}{\mathsf{A}_{m-1,i}^2} \left( (C \varepsilon_{m-1}^{2\delta})^{2N_*} + \varepsilon_{m-1}^{4\delta} \mathcal{F}^4 \right) + \mathbf{1}_{i=1} \left( \frac{C \varepsilon_{m-1}^{2\delta} (C \varepsilon_{m-1}^{2\delta})^{N_*}}{\mathsf{A}_{m-1,1}^2} \mathcal{F}^2 + \frac{C \varepsilon_{m-1}^{4\delta}}{\mathsf{A}_{m-1,1}^2} \mathcal{F}^4 \right) \\
&\quad + \mathbf{1}_{i \geq 2} \left( \frac{C \varepsilon_{m-1}^{4\delta} \mathsf{A}_{m-1,i-2}^2}{\mathsf{A}_{m-1,i}^2} \mathcal{F}^4 + \frac{C \varepsilon_{m-1}^{2\delta} \mathsf{A}_{m-1,i-1}^2}{\mathsf{A}_{m-1,i}^2} \mathcal{F}^2 \right). \tag{4.58}
\end{aligned}$$

We note that upon taking  $N_*$  sufficiently large as in (2.6), and using (4.40), we may ensure that

$$(C \varepsilon_{m-1}^{2\delta})^{N_*} \leq \varepsilon_{m-1}^{2\delta}, \tag{4.59}$$

<sup>13</sup>The inequality  $\mathcal{G} \leq 1/2$  follows from  $\varepsilon_{m-1} \leq 1$  and  $C_0 \geq 2C$ .

so that in (4.58) we may bound  $(C\varepsilon_{m-1}^{2\delta})^{2N_*} \leq \varepsilon_{m-1}^{4\delta} \mathcal{F}^4$ , and  $\varepsilon_{m-1}^{2\delta} (C\varepsilon_{m-1}^{2\delta})^{N_*} \mathcal{F}^2 \leq \varepsilon_{m-1}^{4\delta} \mathcal{F}^4$ .

In order to initiate the induction in  $n \geq 0$ , we first consider estimate (4.58) for  $n = 0$ . Recalling the definition (4.42), and the bound  $\varepsilon_{m-1}^{2\delta} \mathcal{F}^2 \leq C_0^{-1}$  (which is equivalent to (4.40)), (4.58) becomes

$$\begin{aligned} (D_0^{(i)})^2 &\leq \mathbf{1}_{i=1} \frac{C\varepsilon_{m-1}^{4\delta} \mathcal{F}^4}{\mathsf{A}_{m-1,1}^2} + \mathbf{1}_{i \geq 2} \left( \frac{C\mathsf{A}_{m-1,i-1}^2}{\mathsf{A}_{m-1,i}^2} (\varepsilon_{m-1}^{4\delta} \mathcal{F}^4 + \varepsilon_{m-1}^{2\delta} \mathcal{F}^2) + \frac{C\mathsf{A}_{m-1,i-2}^2}{\mathsf{A}_{m-1,i}^2} \varepsilon_{m-1}^{4\delta} \mathcal{F}^4 \right) \\ &\leq \mathbf{1}_{i \in \{1,2\}} \frac{C}{C_0^2} + \mathbf{1}_{i=2} C\varepsilon_{m-1}^{2\delta} \mathcal{F}^2 (1 + \varepsilon_{m-1}^{2\delta} \mathcal{F}^2) + \mathbf{1}_{i \geq 3} \frac{C(1 + \varepsilon_{m-1}^{2\delta} \mathcal{F}^2)}{C_0} \leq \frac{2C}{C_0}. \end{aligned} \quad (4.60)$$

As long as  $C_0$  is taken to be sufficiently large, this established the bound  $D_0^{(i)} \leq 1$ . We now inductively assume  $D_k^{(i)} \leq 1$  for all  $k \in \{0, 1, \dots, n-1\}$ , and aim to establish that  $D_n^{(i)} \leq 1$ ; in turn this would conclude the proof of (4.41). To do so, we return to (4.58), use (4.59), the definition (4.42), and the inductive bound  $D_k^{(i)} \leq 1$  for  $k \leq n-1$ , to conclude

$$\begin{aligned} (D_n^{(i)})^2 &\leq \frac{C}{\mathcal{F}^2} + \frac{2C}{\mathcal{F}^4} + \frac{2C\varepsilon_{m-1}^{2\delta} \mathsf{A}_{m-1,i-1}}{\mathsf{A}_{m-1,i}} \\ &\quad + \mathbf{1}_{i=1} \frac{C\varepsilon_{m-1}^{4\delta}}{\mathsf{A}_{m-1,1}^2} \mathcal{F}^4 + \mathbf{1}_{i \geq 2} \left( \frac{C\mathsf{A}_{m-1,i-1}^2}{\mathsf{A}_{m-1,i}^2} (\varepsilon_{m-1}^{4\delta} \mathcal{F}^4 + \varepsilon_{m-1}^{2\delta} \mathcal{F}^2) + \frac{C\mathsf{A}_{m-1,i-2}^2}{\mathsf{A}_{m-1,i}^2} \varepsilon_{m-1}^{4\delta} \mathcal{F}^4 \right). \end{aligned}$$

The second line of the above estimate precisely matches the upper bound in (4.60), which was shown to be  $\leq 2CC_0^{-1}$  under the standing assumptions. Using that  $\mathcal{F} \geq C_0 \geq 1$ , and recalling the definition (4.41), we may bound also the first line of the above estimate, and finally deduce

$$(D_n^{(i)})^2 \leq \frac{5C}{C_0^2} + \mathbf{1}_{i=1} \frac{2C}{C_0^3} + \mathbf{1}_{i=2} 2C\varepsilon_{m-1}^{2\delta} + \mathbf{1}_{i \geq 3} \frac{2C\varepsilon_{m-1}^{\delta}}{C_0^{3/2}} \leq \frac{5C}{C_0^2} + \mathbf{1}_{i \geq 1} \frac{4C}{C_0^3}.$$

Upon choosing  $C_0$  sufficiently large with respect to  $C$ , we establish the bound necessary for the inductive step  $D_n^{(i)} \leq 1$ , and thus conclude the proof of the Lemma.  $\square$

Direct consequences of the bounds in Lemma 4.2 and of the definition (4.28) are the following regularity estimates for  $T_{m-1}$ .

**Lemma 4.3** (Estimates on  $T_{m-1}$ ). *Let  $C := 4C_0^3$ , where  $C_0 \geq 1$  is the universal constant from Lemma 4.2. If  $\varepsilon_{m-1}$  is small enough to ensure*

$$\varepsilon_{m-1}^{1+\gamma/2} \leq R_{\theta_0} \quad \text{and} \quad \varepsilon_{m-1}^{2\delta} \leq C^{-1}, \quad (4.61)$$

then, for every  $n \in \mathbb{N}_0$  we have

$$\|\nabla^n T_{m-1}\|_{L^\infty([0,1];L^2(\mathbb{T}^2))} + \kappa_{m-1}^{1/2} \|\nabla^{n+1} T_{m-1}\|_{L^2((0,1) \times \mathbb{T}^2)} \leq C_{N_*} \|\theta_0\|_{L^2(\mathbb{T}^2)} n! (C\varepsilon_{m-1}^{-1-\gamma/2})^n, \quad (4.62)$$

where we have defined  $C_{N_*} := 2^{2N_*} (2N_*)!$ .

*Proof.* Assumption (4.61) gives that

$$C_0^3 \varepsilon_{m-1}^{2\delta} (1 \vee \varepsilon_{m-1}^{2+\gamma} R_{\theta_0}^{-2}) = C_0^3 \varepsilon_{m-1}^{2\delta} \leq C_0^3 C^{-1} = 1/4. \quad (4.63)$$

Thus assumption (4.40) holds, and we are allowed to apply Lemma 4.2. For compactness of notation, denote the left side of (4.62) as

$$\mathsf{F}_n := \max_{|\alpha|=n} \left( \|\partial^\alpha T_{m-1}\|_{L^\infty([0,1];L^2(\mathbb{T}^2))} + \kappa_{m-1}^{1/2} \|\nabla \partial^\alpha T_{m-1}\|_{L^2((0,1) \times \mathbb{T}^2)} \right).$$

From (4.28) and the identification  $\theta_{m-1} = V_{m-1}^{(0)}$ , we have that

$$T_{m-1} = \sum_{i=0}^{N_*} V_{m-1}^{(i)}.$$

Hence, the bound (4.41), the definition (4.42), and assumption (4.40) imply that

$$\begin{aligned} \frac{F_n}{n! (C_0 \varepsilon_{m-1}^{-1-\gamma/2} \vee C_0 R_{\theta_0}^{-1})^n} &\leq \|\theta_0\|_{L^2(\mathbb{T}^2)} \sum_{i=0}^{N_*} A_{m-1,i} \frac{(n+2i)!}{n!} \\ &\leq (2N_*)! \|\theta_0\|_{L^2(\mathbb{T}^2)} \sum_{i=0}^{N_*} \binom{n+2i}{n} \frac{(2i)!}{(2N_*)!} \\ &\leq 2^{n+2N_*} (2N_*)! \|\theta_0\|_{L^2(\mathbb{T}^2)}. \end{aligned} \quad (4.64)$$

With assumption (4.61) and the definition  $C = 4C_0^3$ , the proof is completed.  $\square$

In the above proof we have merely used that for all  $i \geq 0$  the amplitude coefficients  $A_{m-1,i}$  appearing in (4.42) satisfy the bound  $A_{m-1,i} \leq 1$ . Now, we use the precise structure of these coefficients to deduce two further consequences.

**Lemma 4.4** ( $T_{m-1}$  and  $\theta_{m-1}$  are close and  $T_{m-1}$  almost solves (4.9)). *Under the assumptions of Lemma 4.3, we have that*

$$\begin{aligned} \|T_{m-1} - \theta_{m-1}\|_{L^\infty((0,1);L^2(\mathbb{T}^2))} + \kappa_{m-1}^{1/2} \|\nabla T_{m-1} - \nabla \theta_{m-1}\|_{L^2((0,1) \times \mathbb{T}^2)} \\ \leq (2N_*)! (C_{(4.62)} \varepsilon_{m-1}^{2\delta}) \|\theta_0\|_{L^2(\mathbb{T}^2)}, \end{aligned} \quad (4.65)$$

and the error term appearing on the right side of (4.13) satisfies

$$\|\nabla^n \mathbf{e}_{m-1}\|_{L^2((0,1) \times \mathbb{T}^2)} \leq CC_{N_*} \kappa_{m-1}^{1/2} n! (C_{(4.62)} \varepsilon_{m-1}^{-1-\gamma/2})^n (C_{(4.62)} \varepsilon_{m-1}^{2\delta})^{N_*/2} \|\theta_0\|_{L^2(\mathbb{T}^2)}, \quad (4.66)$$

for  $n \in \mathbb{N}_0$  and a universal constant  $C \geq 1$ .

*Proof.* In order to prove (4.65), we recall from (4.28) that  $T_{m-1} - \theta_{m-1} = \sum_{i=1}^{N_*} V_{m-1}^{(i)}$ . Therefore, similarly to (4.64) with  $n = 0$ , we deduce from the bound (4.41), the definition (4.42), and assumption (4.61) (which implies (4.40) and also (4.63)), that

$$\begin{aligned} &\|T_{m-1} - \theta_{m-1}\|_{L^\infty((0,1);L^2(\mathbb{T}^2))} + \kappa_{m-1}^{1/2} \|\nabla T_{m-1} - \nabla \theta_{m-1}\|_{L^2((0,1) \times \mathbb{T}^2)} \\ &\leq \|\theta_0\|_{L^2(\mathbb{T}^2)} \sum_{i=1}^{N_*} (2i)! A_{m-1,i} \\ &\leq C_0^3 \varepsilon_{m-1}^{2\delta} \|\theta_0\|_{L^2(\mathbb{T}^2)} \sum_{i=1}^{N_*} (2i)! (C_0^3 \varepsilon_{m-1}^{2\delta})^{\frac{(i-2)_+}{2}} \\ &\leq C_{(4.62)} \varepsilon_{m-1}^{2\delta} \|\theta_0\|_{L^2(\mathbb{T}^2)} \sum_{i=1}^{N_*} (2i)! 2^{-(i-2)_+ - 2} \\ &\leq (2N_*)! C_{(4.62)} \varepsilon_{m-1}^{2\delta} \|\theta_0\|_{L^2(\mathbb{T}^2)}. \end{aligned}$$

This gives (4.65).

In view of the definition of  $\mathbf{e}_{m-1}$  in (4.14), proving (4.66) for  $n \in \{0, 1, 2\}$  amounts to combining the  $V_{m-1}^{(N_*)}$  estimate from (4.41), the  $\mathbf{s}_{m-1}$  bound in (4.39), and the estimate  $|\mathbf{K}_m - \kappa_{m-1}| \leq C\kappa_{m-1}$  (which follows from (3.28)):

$$\begin{aligned} & \|\partial^\alpha \mathbf{e}_{m-1}\|_{L^2((0,1) \times \mathbb{T}^2)} \\ & \leq C\kappa_{m-1} \|\partial^\alpha \nabla V_{m-1}^{(N_*)}\|_{L^2((0,1) \times \mathbb{T}^2)} + C\kappa_{m-1} \sum_{\beta \leq \alpha} \frac{|\alpha|!}{|\beta|!} (C\varepsilon_{m-1}^{-1})^{|\alpha-\beta|} \|\partial^\beta \nabla V_{m-1}^{(N_*)}\|_{L^2((0,1) \times \mathbb{T}^2)} \\ & \leq C\kappa_{m-1}^{1/2} (2N_* + |\alpha|)! A_{m-1, N_*} \|\theta_0\|_{L^2(\mathbb{T}^2)} (C_0 \varepsilon_{m-1}^{-1-\gamma/2})^{|\alpha|} \\ & \leq C\kappa_{m-1}^{1/2} (2N_* + |\alpha|)! (C_{(4.62)} \varepsilon_{m-1}^{2\delta})^{N_*/2} \|\theta_0\|_{L^2(\mathbb{T}^2)} (C_0 \varepsilon_{m-1}^{-1-\gamma/2})^{|\alpha|}. \end{aligned}$$

Since  $(2N_* + |\alpha|)! \leq 2^{|\alpha|+2N_*} (2N_*)! |\alpha|! = C_{N_*} 2^{|\alpha|} |\alpha|!$  and  $2C_0 \leq C_{(4.62)}$ , this concludes the proof of (4.66) and thus of the lemma.  $\square$

The estimates on all space derivatives of  $\nabla T_{m-1}$  obtained in Lemma 4.3 imply, when combined with the available bounds on  $\mathbf{b}_{m-1}$  a control on mixed space-and-material derivatives of  $\nabla T_{m-1}$ .

**Lemma 4.5.** *Under the assumptions of Lemma 4.3, for all  $n, \ell \in \mathbb{N}_0$  with  $\ell \geq 1$  and  $n + 2\ell \leq N_*$ , we have*

$$\|\nabla^n \mathcal{D}_{t,m-1}^\ell \nabla T_{m-1}\|_{L^2([0,1] \times \mathbb{T}^2)} \leq C\varepsilon_{m-1}^{3\delta} \kappa_{m-1}^{-1/2} \|\theta_0\|_{L^2(\mathbb{T}^2)} \varepsilon_{m-1}^{-(1+\gamma/2)n} (\tau'_m)^{-\ell} \quad (4.67)$$

for a sufficiently large constant  $C = C(N_*) \geq 1$ .

*Proof.* In (4.67) we only consider  $\ell \geq 1$  because for  $\ell = 0$  a different bound is already available in (4.62). In order to prove (4.67), we recall from (4.10)–(4.12) that  $T_{m-1} = T_{m-1}^{N_*}$ , where the functions  $\{\nabla T_{m-1}^{(i)}\}_{i=0}^{N_*}$  solve

$$\begin{aligned} \mathcal{D}_{t,m-1} \nabla T_{m-1}^{(i)} &= \kappa_{m-1} \Delta \nabla T_{m-1}^{(i)} - \nabla \mathbf{b}_{m-1} \cdot \nabla T_{m-1}^{(i)} \\ &\quad + \mathbf{1}_{\{i \geq 1\}} \nabla \nabla \cdot (\mathbf{K}_m - \kappa_{m-1} \mathbf{I}_2 + \mathbf{s}_{m-1}) \nabla T_{m-1}^{(i-1)}. \end{aligned} \quad (4.68)$$

We claim that for all  $0 \leq i \leq N_*$ ,  $n \in \mathbb{N}_0$ ,  $\ell \in \mathbb{N}$  with  $n + 2\ell \leq N_*$  we have the bound

$$\kappa_{m-1}^{1/2} \max_{|\alpha|=n} \|\partial^\alpha \mathcal{D}_{t,m-1}^\ell \nabla T_{m-1}^{(i)}\|_{L^2([0,1] \times \mathbb{T}^2)} \leq C\varepsilon_{m-1}^{3\delta} \|\theta_0\|_{L^2(\mathbb{T}^2)} (C\varepsilon_{m-1}^{-1-\gamma/2})^n (\tau'_m)^{-\ell}. \quad (4.69)$$

Specializing (4.69) to the case  $i = N_*$  gives (4.67), upon noting that the factor of  $C^n \leq C^{N_*}$  may be absorbed in the constant  $C_{(4.67)}$ .

We first prove (4.69) for  $\ell = 1$ , as this contains the main idea. The generalization to  $\ell \geq 2$  is a matter of accounting, and the upper bound obtained is allowed to have a large amplitude, by a factor of  $\varepsilon_{m-1}^{-3\delta}$ . Since  $T_{m-1}^{(i)} = \sum_{i'=0}^i V_{m-1}^{(i')}$ , by using that  $A_{m-1, i'} \leq 1$  for all  $i' \geq 0$ , we deduce similarly to (4.62) that for all  $0 \leq i \leq N_*$  and  $n \in \mathbb{N}_0$ , we have

$$\kappa_{m-1}^{1/2} \max_{|\alpha|=n} \|\partial^\alpha \nabla T_{m-1}^{(i)}\|_{L^2((0,1) \times \mathbb{T}^2)} \leq C \|\theta_0\|_{L^2(\mathbb{T}^2)} n! (C\varepsilon_{m-1}^{-1-\gamma/2})^n. \quad (4.70)$$

From (2.72), (3.28), (3.39), (4.39), (4.61), (4.70), and the Leibniz rule, it follows that for all  $0 \leq i \leq N_*$ , we have

$$\begin{aligned}
& \|\theta_0\|_{L^2(\mathbb{T}^2)}^{-1} \kappa_{m-1}^{1/2} \max_{|\alpha|=n} \|\partial^\alpha \mathcal{D}_{t,m-1} \nabla T_{m-1}^{(i)}\|_{L^2([0,1] \times \mathbb{T}^2)} \\
& \leq C \kappa_{m-1} (C \varepsilon_{m-1}^{-1-\gamma/2})^{n+2} (n+2)! + (C \varepsilon_{m-1}^{\beta-2}) n! \sum_{k=0}^n (C \varepsilon_{m-1}^{-1})^{n-k} (C \varepsilon_{m-1}^{-1-\gamma/2})^k \\
& \quad + C \kappa_{m-1} (C \varepsilon_{m-1}^{-1-\gamma/2})^2 (n+2)! \sum_{k=0}^{n+2} (C \varepsilon_{m-1}^{-1})^{n-k+2} (C \varepsilon_{m-1}^{-1-\gamma/2})^{k-2} \\
& \leq C (n+2)! \varepsilon_{m-1}^{\beta-2} (C \varepsilon_{m-1}^{-1-\gamma/2})^n.
\end{aligned}$$

Since  $\varepsilon_{m-1}^{\beta-2} \leq \varepsilon_{m-1}^{3\delta} (\tau'_m)^{-1}$  (see (2.12)–(2.13)), the above estimate gives the proof of (4.69) for  $\ell = 1$ .

From (4.68) it is clear that proving (4.69) for  $\ell \geq 2$ , requires a bound for the space-and-material derivatives of  $\mathbf{s}_{m-1}$  (which we recall was defined in (4.7)). For this purpose, for all  $n, \ell \in \mathbb{N}_0$  with  $n + \ell \leq N_*$  we claim that

$$\|\nabla^n \mathcal{D}_{t,m-1}^\ell \mathbf{s}_{m-1}\|_{L^\infty(\mathbb{R} \times \mathbb{R}^2)} \leq C \kappa_{m-1} \varepsilon_{m-1}^{-n} (\tau'_m)^{-\ell}. \quad (4.71)$$

As usual in such terms, we do not keep track of factorials because the constant  $C$  in (4.71) depends (only) on  $N_*$ . When  $\ell = 0$ , the bound (4.71) follows from (4.39). For  $\ell \geq 1$ , we use the Leibniz rule, (2.13), (2.14), (2.16), (2.29), (2.31), (2.67), (2.96), and (3.28),

$$\begin{aligned}
\|\nabla^n \mathcal{D}_{t,m-1}^\ell \mathbf{s}_{m-1}\|_{L^\infty(\mathbb{R} \times \mathbb{R}^2)} & \leq C \sum_{\ell'=0}^{\ell} \sum_{\ell''=0}^{\ell'} \sum_{l \in \mathbb{Z}} \|\partial_t^{\ell-\ell'} \hat{\xi}_{m,l}\|_{L^\infty(\mathbb{R})} \|\partial_t^{\ell'-\ell''} \mathbf{K}_m\|_{L^\infty(\mathbb{R})} \\
& \quad \times \|\nabla^n \mathcal{D}_{t,m-1}^{\ell''} (\nabla X_{m-1,l} \circ X_{m-1,l}^{-1} - \mathbf{I}_2)\|_{L^\infty(\text{supp } \hat{\xi}_{m,l} \times \mathbb{R}^2)} \\
& \leq C \kappa_{m-1} \sum_{\ell''=0}^{\ell} (\tau'_m)^{\ell''-\ell} \varepsilon_{m-1}^{-n} (\varepsilon_{m-1}^{\beta-2})^{\ell''} \\
& \leq C \kappa_{m-1} \varepsilon_{m-1}^{-n} (\tau'_m)^{-\ell}.
\end{aligned}$$

In the last inequality we have used that  $\tau'_m \varepsilon_{m-1}^{\beta-2} \leq C \varepsilon_{m-1}^{3\delta} \leq 1$ . This concludes the proof of (4.71).

With (4.71) in hand, we return to proving (4.69) for  $\ell \geq 2$ . In view of (4.68), in order to estimate higher order material derivatives of  $\nabla T_{m-1}^{(i)}$ , we need to understand the commutator between  $\mathcal{D}_{t,m}^{\ell-1}$  and  $\nabla^2$ . In this direction, from [BMNV23, Lemma A.12] we recall that

$$[\mathcal{D}_{t,m-1}^{\ell-1}, \nabla^2] f = \sum_{\ell'=1}^{\ell-1} \sum_{\ell''=0}^{\ell'} c_{\ell,\ell',\ell''} (\text{ad } \mathcal{D}_{t,m-1})^{\ell''} (\nabla) (\text{ad } \mathcal{D}_{t,m-1})^{\ell'-\ell''} (\nabla) \mathcal{D}_{t,m-1}^{\ell-1-\ell'} f \quad (4.72)$$

where  $c_{\ell,\ell',\ell''} > 0$  are explicitly computable combinatorial coefficients, and we recall from (2.91) that  $(\text{ad } \mathcal{D}_{t,m-1})^r (\nabla)$  is a first order differential operator for any  $r \geq 0$ . With (4.72), we return to

(4.68) and obtain (ignoring the precise contraction of tensors) that

$$\begin{aligned}
\mathcal{D}_{t,m-1}^\ell \nabla T_{m-1}^{(i)} &= \kappa_{m-1} \sum_{\ell'=0}^{\ell-1} \sum_{\ell''=0}^{\ell'} c_{\ell,\ell',\ell''} (\text{ad } \mathcal{D}_{t,m-1})^{\ell''}(\nabla) (\text{ad } \mathcal{D}_{t,m-1})^{\ell'-\ell''}(\nabla) \mathcal{D}_{t,m-1}^{\ell-1-\ell'} \nabla T_{m-1}^{(i)} \\
&\quad - \sum_{\ell'=0}^{\ell-1} \binom{\ell-1}{\ell'} \mathcal{D}_{t,m-1}^{\ell'} \nabla \mathbf{b}_{m-1} \cdot \mathcal{D}_{t,m-1}^{\ell-1-\ell'} \nabla T_{m-1}^{(i)} \\
&\quad + \mathbf{1}_{\{i \geq 1\}} \sum_{\ell'=0}^{\ell-1} \sum_{\ell''=0}^{\ell'} c_{\ell,\ell',\ell''} (\text{ad } \mathcal{D}_{t,m-1})^{\ell''}(\nabla) (\text{ad } \mathcal{D}_{t,m-1})^{\ell'-\ell''}(\nabla) \\
&\quad \times \mathcal{D}_{t,m-1}^{\ell-1-\ell'} (\mathbf{K}_m - \kappa_{m-1} \mathbf{I}_2 + \mathbf{s}_{m-1}) \nabla T_{m-1}^{(i-1)}. \tag{4.73}
\end{aligned}$$

We note that the number of  $\mathcal{D}_{t,m-1}$  material derivatives acting on  $\nabla T_{m-1}^{(i)}$  and  $\nabla T_{m-1}^{(i-1)}$  on the right side of (4.73) is at most  $\ell-1$ , whereas on the left side of (4.73) we have  $\ell$ -many  $\mathcal{D}_{t,m-1}$  material derivatives acting on  $\nabla T_{m-1}^{(i)}$ . As such the bound (4.69) is established inductively in  $\ell \geq 1$ , with the base step  $\ell=1$  being already proven.

In order to bound the terms on the right side of (4.73), we note that (2.77) implies that

$$\begin{aligned}
&\|\nabla^n (\text{ad } \mathcal{D}_{t,m-1})^r(\nabla) f\|_{L^p(\mathbb{R} \times \mathbb{R}^2)} \\
&\leq C (\varepsilon_{m-1}^{\beta-2})^r \|\nabla^{n+1} f\|_{L^p(\mathbb{R} \times \mathbb{R}^2)} + C (\varepsilon_{m-1}^{-1})^n (\varepsilon_{m-1}^{\beta-2})^r \|\nabla f\|_{L^p(\mathbb{R} \times \mathbb{R}^2)} \tag{4.74}
\end{aligned}$$

holds for all  $n+r \leq N_*$  and  $f \in W^{n+1,p}$ . By combining (4.74) and (4.73), for  $n+2\ell \leq N_*$  with  $\ell \geq 2$ , we obtain that

$$\begin{aligned}
&\|\nabla^n \mathcal{D}_{t,m-1}^\ell \nabla T_{m-1}^{(i)}\|_{L^2([0,1] \times \mathbb{R}^2)} \\
&\leq C \kappa_{m-1} \sum_{\ell'=0}^{\ell-1} \varepsilon_{m-1}^{(\beta-2)\ell'} \left( \|\nabla^{n+2} \mathcal{D}_{t,m-1}^{\ell-1-\ell'} \nabla T_{m-1}^{(i)}\|_{L^2((0,1) \times \mathbb{T}^2)} + \varepsilon_{m-1}^{-(n+1)} \|\nabla \mathcal{D}_{t,m-1}^{\ell-1-\ell'} \nabla T_{m-1}^{(i)}\|_{L^2((0,1) \times \mathbb{T}^2)} \right) \\
&\quad + C \sum_{n'=0}^n \sum_{\ell'=0}^{\ell-1} \|\nabla^{n'} \mathcal{D}_{t,m-1}^{\ell'} \nabla \mathbf{b}_{m-1}\|_{L^\infty((0,1) \times \mathbb{T}^2)} \|\nabla^{n-n'} \mathcal{D}_{t,m-1}^{\ell-1-\ell'} \nabla T_{m-1}^{(i)}\|_{L^2((0,1) \times \mathbb{T}^2)} \\
&\quad + C \mathbf{1}_{\{i \geq 1\}} \sum_{\ell'=0}^{\ell-1} \sum_{\ell''=0}^{\ell-1-\ell'} \varepsilon_{m-1}^{(\beta-2)\ell'} \\
&\quad \times \left( \|\nabla^{n+2} (\mathcal{D}_{t,m-1}^{\ell-1-\ell'-\ell''} (\mathbf{K}_m - \kappa_{m-1} \mathbf{I}_2 + \mathbf{s}_{m-1}) \mathcal{D}_{t,m-1}^{\ell''} \nabla T_{m-1}^{(i-1)})\|_{L^2((0,1) \times \mathbb{T}^2)} \right. \\
&\quad \left. + \varepsilon_{m-1}^{-(n+1)} \|\nabla (\mathcal{D}_{t,m-1}^{\ell-1-\ell'-\ell''} (\mathbf{K}_m - \kappa_{m-1} \mathbf{I}_2 + \mathbf{s}_{m-1}) \mathcal{D}_{t,m-1}^{\ell''} \nabla T_{m-1}^{(i-1)})\|_{L^2((0,1) \times \mathbb{T}^2)} \right). \tag{4.75}
\end{aligned}$$

Recalling the  $\nabla \mathbf{b}_{m-1}$  bound in (2.77), the  $\mathbf{s}_{m-1}$  estimate (4.71), the  $\mathbf{K}_m$  bound in (3.28), the  $\nabla T_{m-1}^{(i)}$  bound with no material derivatives (4.70), the inductive bound (4.69) for  $\mathcal{D}_{t,m-1}^\ell \nabla T_{m-1}^{(i)}$  with  $1 \leq \ell' \leq \ell-1$ , and the parameter inequality  $\varepsilon_{m-1}^{\beta-2} \tau'_m = \varepsilon_{m-1}^{3\delta} \ll 1$ , we obtain from the above

estimate that

$$\begin{aligned}
& \|\theta_0\|_{L^2(\mathbb{T}^2)}^{-1} \kappa_{m-1}^{1/2} \|\nabla^n \mathcal{D}_{t,m-1}^\ell \nabla T_{m-1}^{(i)}\|_{L^2([0,1] \times \mathbb{T}^2)} \\
& \leq C \kappa_{m-1} \sum_{\ell'=0}^{\ell-1} \varepsilon_{m-1}^{(\beta-2)\ell'} \left( (\varepsilon_{m-1}^{-1-\gamma/2})^{n+2} (\tau'_m)^{-(\ell-1-\ell')} + \varepsilon_{m-1}^{-(n+1)} (\varepsilon_{m-1}^{-1-\gamma/2}) (\tau'_m)^{-(\ell-1-\ell')} \right) \\
& \quad + C \varepsilon_{m-1}^{\beta-2} \sum_{n'=0}^n \sum_{\ell'=0}^{\ell-1} (\varepsilon_{m-1}^{\beta-2})^{\ell'} (\varepsilon_{m-1}^{-1})^{n'} (\varepsilon_{m-1}^{-1-\gamma/2})^{n-n'} (\tau'_m)^{-(\ell-1-\ell')} \\
& \quad + C \mathbf{1}_{\{i \geq 1\}} \kappa_{m-1} \sum_{\ell'=0}^{\ell-1} \sum_{\ell''=1}^{\ell-1-\ell'} \sum_{n'=0}^{n+2} \varepsilon_{m-1}^{(\beta-2)\ell'} (\varepsilon_{m-1}^{-1})^{n+2-n'} (\tau'_m)^{-(\ell-1-\ell'-\ell'')} (\varepsilon_{m-1}^{-1-\gamma/2})^{n'} (\tau'_m)^{-\ell''} \\
& \leq C \left( \kappa_{m-1} \varepsilon_{m-1}^{-2-\gamma} + \varepsilon_{m-1}^{\beta-2} \right) (\varepsilon_{m-1}^{-1-\gamma/2})^n \sum_{\ell'=0}^{\ell-1} \varepsilon_{m-1}^{(\beta-2)\ell'} (\tau'_m)^{-(\ell-1-\ell')} \\
& \leq C \varepsilon_{m-1}^{\beta-2} (\varepsilon_{m-1}^{-1-\gamma/2})^n (\tau'_m)^{-(\ell-1)} \\
& = C \varepsilon_{m-1}^{3\delta} (\varepsilon_{m-1}^{-1-\gamma/2})^n (\tau'_m)^{-\ell}
\end{aligned} \tag{4.75}$$

for all  $n+2\ell \leq N_*$ . Note that the bound on the term  $\nabla^{n+2} \mathcal{D}_{t,m-1}^{\ell-1-\ell'} \mathbf{s}_{m-1}$ , cf. (4.71), requires that  $n+2+\ell-1-\ell' \leq N_*$ ; this condition holds because  $n+2+\ell-1-\ell' \leq n+1+\ell \leq n+2\ell$  for  $\ell \geq 1$ . By induction on  $\ell$ , this concludes the proof of (4.69), and thus of the Lemma.  $\square$

#### 4.4 Estimates for $\tilde{H}_m$

Before estimating the function  $\tilde{H}_m$  defined in (4.21), we need to obtain estimates for space-and-material derivatives of the tensors  $\mathbf{A}_{m,n,r}$  defined in (4.20).

**Proposition 4.6.** *Under the assumptions of Lemma 4.3, there exists a constant  $C \geq 1$  such that for all  $0 \leq r \leq N_*/2$ ,  $m \in \mathbb{N}$ , and  $0 \leq n \leq N_* - 1$  we have*

$$\|\nabla^k \mathcal{D}_{t,m-1}^\ell \mathbf{A}_{m,n,r}\|_{L^2([0,1] \times \mathbb{T}^2)} \leq C \|\theta_0\|_{L^2(\mathbb{T}^2)} (\varepsilon_m^2 \kappa_m^{-1}) \kappa_{m-1}^{-1/2} \varepsilon_{m-1}^{-(1+\gamma/2)k} (\tau'_m)^{-\ell-r} \tag{4.76}$$

for all  $k+2\ell \leq N_* - 2r$ .

*Proof.* We appeal to the space-and-material bounds for  $\nabla \mathbf{b}_{m-1}$  available from (2.76), the space-and-material estimates for  $\nabla X_{m-1,\ell} \circ X_{m-1,\ell}^{-1}$  in (2.96), the space-and-material bounds for  $\nabla T_{m-1}$  in (4.62) and (4.67), the time derivative bounds for  $L_{m,n}^{\kappa_m}$  in (3.24), and the time derivative bounds for  $\hat{\xi}_{m,l}$  in (2.31). Using these bounds, the product rule, and the definition of  $\mathbf{A}_{m,n,0}$  in (4.20), we deduce

$$\|\nabla^k \mathcal{D}_{t,m-1}^\ell \mathbf{A}_{m,n,0}\|_{L^2([0,1] \times \mathbb{T}^2)} \leq C \|\theta_0\|_{L^2(\mathbb{T}^2)} (\varepsilon_m^2 \kappa_m^{-1}) \kappa_{m-1}^{-1/2} \varepsilon_{m-1}^{-(1+\gamma/2)k} (\tau'_m)^{-\ell} \tag{4.77}$$

for all  $k+2\ell \leq N_*$ . Here we have used implicitly the bounds  $\varepsilon_{m-1}^{\beta-2} \leq C(\tau'_m)^{-1}$  and  $\kappa_m \varepsilon_m^{-2} \leq C(\tau'_m)^{-1}$ . Inductively in  $r$ , it is then direct to establish

$$\|\nabla^k \mathcal{D}_{t,m-1}^\ell \mathbf{A}_{m,n,r}\|_{L^2([0,1] \times \mathbb{T}^2)} \leq C \|\theta_0\|_{L^2(\mathbb{T}^2)} (\varepsilon_m^2 \kappa_m^{-1}) \kappa_{m-1}^{-1/2} \varepsilon_{m-1}^{-(1+\gamma/2)k} (\tau'_m)^{-\ell-r} \tag{4.78}$$

but only for  $k$  and  $\ell$  that satisfy  $k+2(\ell+r) \leq N_*$ . To see this, note that the recursion relation in (4.20) gives  $\mathbf{A}_{m,n,r+1} = \mathcal{D}_{t,m-1} \mathbf{A}_{m,n,r} + \nabla \mathbf{b}_{m-1} \mathbf{A}_{m,n,r}$ , with suitable contraction. If only the first term in this relation would be present, then (4.78) would simply follow by induction. The second

term in this relation requires that we use the Leibniz rule to decompose  $\nabla^k \mathcal{D}_{t,m-1}^\ell(\nabla \mathbf{b}_{m-1} \mathbf{A}_{m,n,r}) = \sum_{k'=0}^k \sum_{\ell'=0}^\ell \binom{k}{k'} \binom{\ell}{\ell'} \nabla^{k-k'} \mathcal{D}_{t,m-1}^{\ell-\ell'} \nabla \mathbf{b}_{m-1} \nabla^{k'} \mathcal{D}_{t,m-1}^{\ell'} \mathbf{A}_{m,n,r}$ . The desired bound at level  $r+1$  is then a consequence of (2.76) and (4.78) at level  $r$ .  $\square$

**Proposition 4.7.** *Under the assumptions of Lemma 4.3, there exists a constant  $C > 0$ , which only depends only on  $N_*$ , such that*

$$\|\tilde{H}_m(t, \cdot)\|_{L_t^\infty L_x^2([0,1] \times \mathbb{T}^2)} \leq C \varepsilon_{m-1}^\delta \|\theta_0\|_{L^2(\mathbb{T}^2)}. \quad (4.79)$$

and

$$\|\nabla \tilde{H}_m(t, \cdot)\|_{L_t^2 L_x^2([0,1] \times \mathbb{T}^2)} \leq C \varepsilon_{m-1}^{4\delta} \kappa_{m-1}^{-1/2} \|\theta_0\|_{L^2(\mathbb{T}^2)}. \quad (4.80)$$

*Proof.* Recall that

$$\tilde{H}_m(t, x) = \nabla \cdot \sum_{r=0}^{N_*/2} \sum_{n=0}^{N_*-1} \mathbf{A}_{m,n,r}(t, x) \mathbf{q}_{m,n,r+1}(t). \quad (4.81)$$

*Step 1.* The uniform-in-time estimate (4.79). First, we observed that for all  $l \in \mathbb{Z}$ , and  $t \in [(l-1/2)\tau_m'', (l+1/2)\tau_m'']$ , since the flow  $X_{m-1,l}(t, \cdot)$  is volume preserving, we have that  $\|\tilde{H}_m(t, \cdot)\|_{L_x^2} = \|\tilde{H}_m \circ X_{m-1,l}(t, \cdot)\|_{L_x^2}$ . Second, we note that by the construction of  $\hat{\zeta}_{m,l}$ , this function vanishes identically in a  $2\tau_m'$ -neighborhood of  $(l \pm 1/2)\tau_m''$  (see (2.29)), and thus by the definition of  $L_{m,n}^{\kappa_m}$  in (3.22), we have

$$(\partial_t^\ell L_{m,n}^{\kappa_m})((l \pm 1/2)\tau_m'') = 0 \quad (4.82)$$

for all  $\ell \in \mathbb{N}$ . In turn, (4.82) and the recursive definition of the  $\mathbf{A}_{m,n,r}$  tensors in (4.20) gives that  $\mathbf{A}_{m,n,r}((l \pm 1/2)\tau_m'') = 0$  for all  $0 \leq r \leq N_*/2$ , and thus

$$\tilde{H}_m((l \pm 1/2)\tau_m'') = 0,$$

for all  $l \in \mathbb{Z}$ . Combining these two observations with the fundamental theorem of calculus in time, we deduce that for all  $t \in [(l-1/2)\tau_m'', (l+1/2)\tau_m'']$

$$\begin{aligned} \|\tilde{H}_m(t, \cdot)\|_{L_x^2}^2 &= \int_{(l-1/2)\tau_m''}^t \frac{d}{dt'} \|\tilde{H}_m(t', X_{m-1,l}(t', \cdot))\|_{L_x^2}^2 dt' \\ &= 2 \int_{(l-1/2)\tau_m''}^t \int_{\mathbb{T}^2} \tilde{H}_m(t', X_{m-1,l}(t', x)) (\mathcal{D}_{t,m-1} \tilde{H}_m)(t', X_{m-1,l}(t', x)) dx dt' \\ &\leq 2(\tau_m'')^{1/2} \|\mathcal{D}_{t,m-1} \tilde{H}_m\|_{L^2([0,1] \times \mathbb{T}^2)} \sup_{t \in [(l-1/2)\tau_m'', (l+1/2)\tau_m'']} \|\tilde{H}_m(t, \cdot)\|_{L_x^2}. \end{aligned} \quad (4.83)$$

Next, using (4.81), (3.31) and (4.76), we derive

$$\begin{aligned} &\|\mathcal{D}_{t,m-1} \tilde{H}_m\|_{L^2([0,1] \times \mathbb{T}^2)} \\ &\leq \sum_{r=0}^{N_*/2} \sum_{n=0}^{N_*-1} \left( \|\nabla \mathcal{D}_{t,m-1} \mathbf{A}_{m,n,r}\|_{L^2([0,1] \times \mathbb{T}^2)} \|\mathbf{q}_{m,n,r+1}\|_{L^\infty([0,1])} \right. \\ &\quad + \|\nabla \mathbf{b}_{m-1}\|_{L^\infty([0,1] \times \mathbb{T}^2)} \|\nabla \mathbf{A}_{m,n,r}\|_{L^2([0,1] \times \mathbb{T}^2)} \|\mathbf{q}_{m,n,r+1}\|_{L^\infty([0,1])} \\ &\quad \left. + \|\nabla \mathbf{A}_{m,n,r}\|_{L^2([0,1] \times \mathbb{T}^2)} \|\partial_t \mathbf{q}_{m,n,r+1}\|_{L^\infty([0,1])} \right) \\ &\leq C \|\theta_0\|_{L^2(\mathbb{T}^2)} \sum_{r=0}^{N_*/2} \sum_{n=0}^{N_*-1} \left( \left\{ (\varepsilon_m^2 \kappa_m^{-1}) \kappa_{m-1}^{-1/2} \varepsilon_{m-1}^{-(1+\gamma/2)} (\tau_m')^{-r-1} \right\} \left\{ \frac{a_m^2 \varepsilon_m^2}{n!} \left( \frac{C \varepsilon_m^2}{\kappa_m \tau_m} \right)^n \frac{(C \tau_m)^{r+1}}{(r+1)!} \right\} \right. \\ &\quad \left. + \left\{ (\varepsilon_m^2 \kappa_m^{-1}) \kappa_{m-1}^{-1/2} \varepsilon_{m-1}^{-(1+\gamma/2)} (\tau_m')^{-r} \right\} \left\{ \frac{a_m^2 \varepsilon_m^2}{n!} \left( \frac{C \varepsilon_m^2}{\kappa_m \tau_m} \right)^n \frac{(C \tau_m)^r}{r!} \right\} \right). \end{aligned}$$

We next use that

$$\sum_{r=0}^{\infty} \left( \frac{C\tau_m}{\tau'_m} \right)^r \leq \sum_{r=0}^{\infty} (C\varepsilon_{m-1}^{\delta})^r \leq 2, \quad (4.84)$$

and

$$\sum_{n=0}^{\infty} \left( \frac{C\varepsilon_m^2}{\kappa_m \tau_m} \right)^n \leq \sum_{n=0}^{\infty} (C\varepsilon_{m-1}^{2\delta})^n \leq 2. \quad (4.85)$$

Combining the three displays above, we arrive at

$$\|\mathcal{D}_{t,m-1} \tilde{H}_m\|_{L^2([0,1] \times \mathbb{T}^2)} \leq C\|\theta_0\|_{L^2(\mathbb{T}^2)} (a_m^2 \varepsilon_m^4 \kappa_m^{-1}) \kappa_{m-1}^{-1/2} \varepsilon_{m-1}^{-(1+\gamma/2)}. \quad (4.86)$$

Returning to (4.83), we take the supremum in time over  $t \in [(l - 1/2)\tau_m'', (l + 1/2)\tau_m'']$ , absorb the suitable term in the left side, and then taking a supremum over  $l \in \mathbb{Z}$ , we deduce

$$\begin{aligned} \|\tilde{H}_m\|_{L^\infty([0,1]; L^2(\mathbb{T}^2))} &\leq C\|\theta_0\|_{L^2(\mathbb{T}^2)} (a_m^2 \varepsilon_m^4 \kappa_m^{-1}) (\tau_m'')^{1/2} \kappa_{m-1}^{-1/2} \varepsilon_{m-1}^{-(1+\gamma/2)} \\ &\leq C\|\theta_0\|_{L^2(\mathbb{T}^2)} \kappa_{m-1} (a_{m-1}^{-1} \varepsilon_{m-1}^{2\delta})^{1/2} \kappa_{m-1}^{-1/2} \varepsilon_{m-1}^{-(1+\gamma/2)} \\ &\leq C\|\theta_0\|_{L^2(\mathbb{T}^2)} \varepsilon_{m-1}^{\delta}. \end{aligned}$$

This concludes the proof of (4.79).

*Step 2.* The gradient estimate (4.80). Using (4.81), (3.31) and (4.76), we obtain

$$\begin{aligned} &\|\nabla \tilde{H}_m\|_{L^2([0,1] \times \mathbb{T}^2)} \\ &\leq \sum_{r=0}^{N_{*}/2} \sum_{n=0}^{N_{*}-1} \|\nabla^2 \mathbf{A}_{m,n,r}\|_{L^2([0,1] \times \mathbb{T}^2)} \|\mathbf{q}_{m,n,r+1}\|_{L^\infty([0,1])} \\ &\leq C\|\theta_0\|_{L^2(\mathbb{T}^2)} \sum_{r=0}^{N_{*}/2} \sum_{n=0}^{N_{*}-1} \left\{ (\varepsilon_m^2 \kappa_m^{-1}) \kappa_{m-1}^{-1/2} \varepsilon_{m-1}^{-2(1+\gamma/2)} (\tau'_m)^{-r} \right\} \left\{ \frac{a_m^2 \varepsilon_m^2}{n!} \left( \frac{C\varepsilon_m^2}{\kappa_m \tau_m} \right)^n \frac{(C\tau_m)^{r+1}}{(r+1)!} \right\}. \end{aligned} \quad (4.87)$$

Inserting the bounds (4.84) and (4.85) into (4.87), we obtain that

$$\begin{aligned} \|\nabla \tilde{H}_m\|_{L^2([0,1] \times \mathbb{T}^2)} &\leq C\|\theta_0\|_{L^2(\mathbb{T}^2)} \left\{ (\varepsilon_m^2 \kappa_m^{-1}) \kappa_{m-1}^{-1/2} \varepsilon_{m-1}^{-2(1+\gamma/2)} \right\} \left\{ a_m^2 \varepsilon_m^2 \tau_m \right\} \\ &= C\|\theta_0\|_{L^2(\mathbb{T}^2)} \frac{a_m^2 \varepsilon_m^4}{\kappa_m} \tau_m \varepsilon_{m-1}^{-(2+\gamma)} \kappa_{m-1}^{-1/2} \\ &= C\|\theta_0\|_{L^2(\mathbb{T}^2)} a_{m-1} \tau_m \kappa_{m-1}^{-1/2} = C\|\theta_0\|_{L^2(\mathbb{T}^2)} \varepsilon_{m-1}^{4\delta} \kappa_{m-1}^{-1/2}, \end{aligned}$$

which proves (4.80).  $\square$

**Proposition 4.8.** *Under the assumptions of Lemma 4.3, there exists a constant  $C > 0$ , which only depends only on  $N_*$ , such that*

$$\|\mathbf{d}_m\|_{L^2([0,1] \times \mathbb{T}^2)} \leq C\kappa_{m-1}^{1/2} (C\varepsilon_{m-1}^{\delta})^{N_{*}/2} \|\theta_0\|_{L^2(\mathbb{T}^2)}. \quad (4.88)$$

*Proof.* Recalling the definition of  $\mathbf{d}_m$  in (4.23) we have

$$\begin{aligned} \|\mathbf{d}_m\|_{L^2([0,1] \times \mathbb{T}^2)} &\leq 2\|\hat{\mathbf{J}}_m - \mathbf{J}_m\|_{L^\infty([0,1])} \sup_{l \in \mathbb{Z}} \|\hat{\xi}_{m,l}(\nabla X_{m-1,l} \circ X_{m-1,l}^{-1})\|_{L^\infty([0,1] \times \mathbb{T}^2)} \|\nabla T_{m-1}\|_{L^2([0,1] \times \mathbb{T}^2)} \\ &\quad + \sum_{n=0}^{N_{*}-1} \|\mathbf{A}_{m,n,N_{*}/2}\|_{L^2([0,1] \times \mathbb{T}^2)} \|\mathbf{q}_{m,n,N_{*}/2}^{\kappa_m}\|_{L^\infty([0,1] \times \mathbb{T}^2)}. \end{aligned}$$

Appealing to the closeness of  $\hat{\mathbf{J}}$  to  $\mathbf{J}$  in (3.25), the  $\mathbf{A}$  bound in (4.76), the  $\mathbf{q}$  bound in (3.31), the  $\nabla T_{m-1}$  estimate in (4.62), the summability in  $n$  from (4.85), and the flow bound (2.96), we deduce

$$\begin{aligned} \|\mathbf{d}_m\|_{L^2([0,1]\times\mathbb{T}^2)} &\leq \frac{Ca_m^2\varepsilon_m^4}{\kappa_m} \left( \frac{\varepsilon_m^2}{\kappa_m\tau_m} \right)^{N_*} \kappa_{m-1}^{-1/2} \|\theta_0\|_{L^2(\mathbb{T}^2)} \\ &\quad + C\|\theta_0\|_{L^2(\mathbb{T}^2)} \sum_{n=0}^{N_*-1} (\varepsilon_m^2\kappa_m^{-1}) \kappa_{m-1}^{-1/2} (\tau'_m)^{-N_*/2} \frac{a_m^2\varepsilon_m^2}{n!} \left( \frac{C\varepsilon_m^2}{\kappa_m\tau_m} \right)^n \tau_m^{N_*/2} \\ &\leq C\|\theta_0\|_{L^2(\mathbb{T}^2)} \kappa_{m-1}^{1/2} (C\varepsilon_{m-1}^{2\delta})^{N_*} + C\|\theta_0\|_{L^2(\mathbb{T}^2)} \kappa_{m-1}^{1/2} (C\varepsilon_{m-1}^\delta)^{N_*/2}. \end{aligned}$$

Since the second of the above two terms is larger, this gives (4.88).  $\square$

## 5. Homogenization cascade up the inertial-convection subrange

In this section, we give the proof of Theorem 1.1. We begin by plugging the ansatz into the advection-diffusion operator and computing the error. This is the purpose of the next subsection. In Section 5.2 we estimate the error term, which is then used in Section 5.3 to complete the main induction step, summarized in Proposition 5.2. The proof of the theorem appears finally in Section 5.4.

### 5.1 Computing the error in the multiscale ansatz

In this subsection, we compute an explicit expression for the error obtained when we insert the two-scale ansatz  $\tilde{\theta}_m$  it into the left side of (4.8). That is, we compute  $(\partial_t - \kappa_m \Delta + \mathbf{b}_m \cdot \nabla) \tilde{\theta}_m$ . The main result is the Big Display on Page 73.

Throughout, we use the abbreviated notations

$$\begin{cases} \tilde{\psi}_{m,k}(t, x) := (\psi_{m,k} \circ X_{m-1,k}^{-1})(t, x), \\ \tilde{\psi}_m(t, x) := \sum_{k \in \mathbb{Z}} \hat{\zeta}_{m,l_k} \zeta_{m,k}(t) \tilde{\psi}_{m,k}(t, x). \end{cases}$$

As in Section 4, we use the notational convention that function compositions are with respect to the space variables only. Observe that the recurrence in (2.36) may be written as

$$\phi_m(t, x) - \phi_{m-1}(t, x) = \tilde{\psi}_m(t, x), \quad \forall m \in \mathbb{N} \cap [1, \infty). \quad (5.1)$$

We proceed by splitting the operator:

$$\begin{aligned} (\partial_t - \kappa_m \Delta + \mathbf{b}_m \cdot \nabla) \tilde{\theta}_m &= (\partial_t + \mathbf{b}_{m-1} \cdot \nabla) \tilde{\theta}_m + (-\kappa_m \Delta + (\mathbf{b}_m - \mathbf{b}_{m-1}) \cdot \nabla) \tilde{\theta}_m \\ &= \underbrace{(\partial_t + \mathbf{b}_{m-1} \cdot \nabla) \tilde{\theta}_m}_{\text{the transport term}} - \underbrace{\nabla \cdot (\kappa_m \mathbf{I}_2 + \tilde{\psi}_m \sigma) \nabla \tilde{\theta}_m}_{\text{the diffusion term}}. \end{aligned} \quad (5.2)$$

We will compute the transport term and the diffusion term separately.

*5.1.1. Computation of the transport term.* It should come as no surprise that we will use Lagrangian coordinates to compute the transport term. We make use of the following two identities:

$$(\partial_t + \mathbf{b}_{m-1} \cdot \nabla) \tilde{\chi}_{m,k} = \partial_t \chi_{m,k} \circ X_{m-1,l_k}^{-1}, \quad (5.3)$$

and

$$\begin{aligned} & (\partial_t + \mathbf{b}_{m-1} \cdot \nabla) (\nabla(T_{m-1} \circ X_{m-1,l}) \circ X_{m-1,l}^{-1}) \\ &= (\nabla X_{m-1,l} \circ X_{m-1,l}^{-1}) \nabla \nabla \cdot ((\mathbf{K}_m + \mathbf{s}_{m-1}) \nabla T_{m-1} + \mathbf{e}_{m-1}). \end{aligned} \quad (5.4)$$

To prove these, recall that if  $Z(t, x)$  is a flow for  $\mathbf{b}_{m-1}$ , that is, a solution of the ODE

$$\partial_t Z = \mathbf{b}_{m-1}(t, Z),$$

then the inverse flow  $Z^{-1}$  satisfies the transport equation (cf. (B.23))

$$(\partial_t + \mathbf{b}_{m-1} \cdot \nabla) Z^{-1} = 0.$$

Moreover, if  $Z^{-1}$  is smooth, then any function of  $Z^{-1}$  also satisfies the same transport equation. Applying this to  $Z = X_{m-1,k}$ , while keeping in mind the convention that function compositions in our formulas are with respect to the spatial variable only, we find that, for any function  $F(t, x)$  of both  $t$  and  $x$ , we have

$$(\partial_t + \mathbf{b}_{m-1} \cdot \nabla) (F \circ Z^{-1}) = \partial_t F \circ Z^{-1}, \quad (5.5)$$

and

$$((\partial_t + \mathbf{b}_{m-1} \cdot \nabla) F) \circ Z = \partial_t (F \circ Z). \quad (5.6)$$

The first claimed identity (5.3) is then immediate from (5.5). To obtain (5.4), we use (5.5), (5.6) and the equation (4.13) for  $T_{m-1}$  as follows:

$$\begin{aligned} & (\partial_t + \mathbf{b}_{m-1} \cdot \nabla) (\nabla(T_{m-1} \circ X_{m-1,l}) \circ X_{m-1,l}^{-1}) \\ &= (\partial_t \nabla(T_{m-1} \circ X_{m-1,l})) \circ X_{m-1,l}^{-1} \\ &= (\nabla \partial_t(T_{m-1} \circ X_{m-1,l})) \circ X_{m-1,l}^{-1} \\ &= (\nabla((\partial_t T_{m-1} + \mathbf{b}_{m-1} \cdot \nabla T_{m-1}) \circ X_{m-1,l})) \circ X_{m-1,l}^{-1} \\ &= (\nabla X_{m-1,l} \circ X_{m-1,l}^{-1}) \nabla [\nabla \cdot ((\mathbf{K}_m + \mathbf{s}_{m-1}) \nabla T_{m-1} + \mathbf{e}_{m-1})]. \end{aligned}$$

We are ready to apply  $(\partial_t + \mathbf{b}_{m-1} \cdot \nabla)$  to both sides of (4.25). Using the product rule, the equation for  $\tilde{H}_m$  in (4.22), the definition of  $\tilde{G}_m$  in (4.18), and the above identities (5.3) and (5.4), we obtain

$$\begin{aligned} (\partial_t + \mathbf{b}_{m-1} \cdot \nabla) \tilde{\theta}_m &= (\partial_t + \mathbf{b}_{m-1} \cdot \nabla) T_{m-1} + \sum_{k \in 2\mathbb{Z}+1} (\partial_t \xi_{m,k} \tilde{\chi}_{m,k}) \cdot \nabla (T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1} \\ &+ \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} (\partial_t \chi_{m,k} \circ X_{m-1,l_k}^{-1}) \cdot \nabla (T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1} \\ &+ \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} \tilde{\chi}_{m,k} \cdot (\nabla X_{m-1,l_k} \circ X_{m-1,l_k}^{-1}) \nabla \nabla \cdot ((\mathbf{K}_m + \mathbf{s}_{m-1}) \nabla T_{m-1} + \mathbf{e}_{m-1}) \\ &+ \nabla \cdot \sum_{l \in \mathbb{Z}} \hat{\xi}_{m,l} (\mathbf{J}_m - \mathbf{K}_m) \nabla (T_{m-1} \circ X_{m-1,l}) \circ X_{m-1,l}^{-1} + \nabla \cdot \mathbf{d}_m. \end{aligned} \quad (5.7)$$

In view of the definition of  $\mathbf{s}_{m-1}$  in (4.7), we can write the equation for  $T_{m-1}$  as

$$(\partial_t + \mathbf{b}_{m-1} \cdot \nabla) T_{m-1} = \nabla \cdot \left( \sum_{l \in \mathbb{Z}} \hat{\xi}_{m,l} \mathbf{K}_m \nabla (T_{m-1} \circ X_{m-1,l}) \circ X_{m-1,l}^{-1} + \mathbf{e}_{m-1} \right).$$

Using this, we can cancel the first term on the right of (5.7) with part of the first term on the last line (the expression involving  $\mathbf{K}_m$ ). Note that we are using here the fact that  $\mathbf{K}_m$  is a scalar matrix, and it therefore commutes with  $(\nabla X_{m-1,l_k} \circ X_{m-1,l_k}^{-1})$ . We therefore obtain

$$\begin{aligned} (\partial_t + \mathbf{b}_{m-1} \cdot \nabla) \tilde{\theta}_m &= \sum_{k \in 2\mathbb{Z}+1} (\partial_t \xi_{m,k} \tilde{\chi}_{m,k}) \cdot \nabla (T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1} \\ &\quad + \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} (\partial_t \chi_{m,k} \circ X_{m-1,l_k}^{-1}) \cdot \nabla (T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1} \\ &\quad + \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} \tilde{\chi}_{m,k} \cdot (\nabla X_{m-1,l_k} \circ X_{m-1,l_k}^{-1}) \nabla \nabla \cdot ((\mathbf{K}_m + \mathbf{s}_{m-1}) \nabla T_{m-1} + \mathbf{e}_{m-1}) \\ &\quad + \mathbf{J}_m : \sum_{l \in \mathbb{Z}} \hat{\xi}_{m,l} \nabla (\nabla (T_{m-1} \circ X_{m-1,l}) \circ X_{m-1,l}^{-1}) + \nabla \cdot (\mathbf{d}_m + \mathbf{e}_{m-1}). \end{aligned} \quad (5.8)$$

Below we will insert the identity (5.8) for the transport term back into the right side of (5.2). The first, third and fifth terms on the right side of (5.8) are “acceptable errors,” that is, we will eventually show that they are negligible for our purposes. The second and fourth terms will cancel some expressions arising in our computation of the diffusion term, which we pursue next.

*5.1.2. Computation of the diffusion term.* We write the diffusive term in divergence form as

$$(-\kappa_m \Delta + (\mathbf{b}_m - \mathbf{b}_{m-1}) \cdot \nabla) \tilde{\theta}_m = -\nabla \cdot ((\kappa_m \mathbf{I}_2 + \tilde{\psi}_m \sigma) \nabla \tilde{\theta}_m). \quad (5.9)$$

Returning to the formula (4.25) to compute the gradient of  $\tilde{\theta}_m$ , we find

$$\begin{aligned} \nabla \tilde{\theta}_m &= \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} (\mathbf{I}_2 + \nabla \tilde{\chi}_{m,k}) \sum_{l \in \mathbb{Z}} \hat{\xi}_{m,l} \nabla (T_{m-1} \circ X_{m-1,l}) \circ X_{m-1,l}^{-1} \\ &\quad + \sum_{l \in \mathbb{Z}} \hat{\xi}_{m,l} (\mathbf{I}_2 - \nabla X_{m-1,l} \circ X_{m-1,l}^{-1}) \nabla T_{m-1} \\ &\quad + \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} \tilde{\chi}_{m,k} \nabla (\nabla (T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1}) + \nabla \tilde{H}_m \end{aligned} \quad (5.10)$$

Inserting this into the operator on the right side of (5.9), we obtain

$$\begin{aligned} &-\nabla \cdot ((\kappa_m \mathbf{I}_2 + \tilde{\psi}_m \sigma) \nabla \tilde{\theta}_m) \\ &= -\nabla \cdot \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} (\kappa_m \mathbf{I}_2 + \tilde{\psi}_m \sigma) (\mathbf{I}_2 + \nabla \tilde{\chi}_{m,k}) \sum_{l \in \mathbb{Z}} \hat{\xi}_{m,l} \nabla (T_{m-1} \circ X_{m-1,l}) \circ X_{m-1,l}^{-1} \\ &\quad - \nabla \cdot \sum_{l \in \mathbb{Z}} \hat{\xi}_{m,l} (\kappa_m \mathbf{I}_2 + \tilde{\psi}_m \sigma) (\mathbf{I}_2 - \nabla X_{m-1,l} \circ X_{m-1,l}^{-1}) \nabla T_{m-1} \\ &\quad - \nabla \cdot \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} (\kappa_m \mathbf{I}_2 + \tilde{\psi}_m \sigma) \tilde{\chi}_{m,k} \nabla (\nabla (T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1}) \\ &\quad - \nabla \cdot (\kappa_m \mathbf{I}_2 + \tilde{\psi}_m \sigma) \nabla \tilde{H}_m. \end{aligned} \quad (5.11)$$

The second, third and fourth terms on the right side of (5.11) are acceptable errors: we will show in the next subsection that they are negligible for our purposes.

Let’s look more closely at what is inside the divergence in the first term on the right side of (5.11). Using the properties (2.25) and (2.26) of the cutoff functions and the fact that both  $\tilde{\chi}_{m,k}^\kappa$  and  $\tilde{\psi}_{m,k}$  vanish when  $k$  is even, we see that

$$\xi_{m,k} \zeta_{m,k'} \tilde{\psi}_{m,k'} = \zeta_{m,k} \tilde{\psi}_{m,k} \mathbf{1}_{\{k=k'\}} \quad \text{and} \quad \xi_{m,k} \zeta_{m,k'} \nabla \tilde{\chi}_{m,k'} = \zeta_{m,k} \nabla \tilde{\chi}_{m,k} \mathbf{1}_{\{k=k'\}}.$$

We therefore obtain

$$\begin{aligned}
& \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} (\kappa_m \mathbf{I}_2 + \tilde{\psi}_m \sigma) (\mathbf{I}_2 + \nabla \tilde{\chi}_{m,k}) \sum_{l \in \mathbb{Z}} \hat{\xi}_{m,l} \nabla (T_{m-1} \circ X_{m-1,l}) \circ X_{m-1,l}^{-1} \\
&= \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} \left( \kappa_m \mathbf{I}_2 + \sum_{k' \in \mathbb{Z}} \hat{\zeta}_{m,l_k} \zeta_{m,k'} \tilde{\psi}_{m,k'} \sigma \right) (\mathbf{I}_2 + \nabla \tilde{\chi}_{m,k}) \sum_{l \in \mathbb{Z}} \hat{\xi}_{m,l} \nabla (T_{m-1} \circ X_{m-1,l}) \circ X_{m-1,l}^{-1} \\
&= \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} (\kappa_m \mathbf{I}_2 + \hat{\zeta}_{m,l_k} \zeta_{m,k} \tilde{\psi}_{m,k} \sigma) (\mathbf{I}_2 + \nabla \tilde{\chi}_{m,k}) \sum_{l \in \mathbb{Z}} \hat{\xi}_{m,l} \nabla (T_{m-1} \circ X_{m-1,l}) \circ X_{m-1,l}^{-1}. \quad (5.12)
\end{aligned}$$

Before we compute the divergence of this expression, we need to use a special property of the shear flow structure, which is that

$$\nabla \cdot (\tilde{\psi}_{m,k} \sigma \nabla \tilde{\chi}_{m,k}) = \nabla \tilde{\psi}_{m,k} \cdot \sigma \nabla \tilde{\chi}_{m,k} = 0. \quad (5.13)$$

Indeed, from (2.19)–(2.20), (3.2) and (3.8), we see that  $\psi_{m,k}$  and  $\chi_{m,k}$  depend only on one coordinate  $x_i$  for some  $i \in \{1, 2\}$ , which is the same for both functions and depends only on  $k$ . Using coordinates with the summation convention, we therefore compute

$$\begin{aligned}
\nabla \tilde{\psi}_{m,k} \cdot \sigma \nabla \tilde{\chi}_{m,k} &= \partial_{x_i} \tilde{\psi}_{m,k} \sigma_{ij} \partial_{x_j} \tilde{\chi}_{m,k} \\
&= (\partial_{x_l} \psi_{m,k} \circ X_{m-1}^{-1})_l \sigma_{ij} (\partial_{x_{l'}} \chi_{m,k} \circ X_{m-1}^{-1})_l \partial_{x_j} (X_{m-1}^{-1})_{l'} \\
&= \underbrace{\sigma_{ij} \partial_{x_i} (X_{m-1}^{-1})_l \partial_{x_j} (X_{m-1}^{-1})_l}_{=0} (\partial_{x_l} \psi \circ X_{m-1}^{-1})_l (\partial_{x_l} \chi_{m,k} \circ X_{m-1}^{-1})_l = 0.
\end{aligned}$$

(See also Remark 3.1.) Using (5.13), we may write

$$\begin{aligned}
& \nabla \cdot ((\kappa_m \mathbf{I}_2 + \hat{\zeta}_{m,l_k} \zeta_{m,k} \tilde{\psi}_{m,k} \sigma) (\mathbf{I}_2 + \nabla \tilde{\chi}_{m,k})) \\
&= \nabla \cdot (\hat{\zeta}_{m,l_k} \zeta_{m,k} \tilde{\psi}_{m,k} \sigma + \kappa_m \nabla \tilde{\chi}_{m,k}) \\
&= \nabla \cdot \left( \hat{\zeta}_{m,l_k} \zeta_{m,k} \tilde{\psi}_{m,k} \sigma + \kappa_m \nabla \chi_{m,k} \circ X_{m-1,l_k}^{-1} + \kappa_m (\nabla X_{m-1,l_k}^{-1} - \mathbf{I}_2) \nabla \chi_{m,k} \circ X_{m-1,l_k}^{-1} \right).
\end{aligned}$$

To proceed, we next use the fact that, for any smooth vector field  $\mathbf{g}$  and smooth, measure-preserving map  $M : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ , we have

$$\nabla \cdot ((\nabla M^{-1} \mathbf{g}) \circ M) = (\nabla \cdot \mathbf{g}) \circ M. \quad (5.14)$$

In view of (5.14), applied with  $\mathbf{g} = \hat{\zeta}_{m,l_k} \zeta_{m,k} \psi_{m,k} \sigma + \kappa_m \nabla \chi_{m,k}$  and  $M = X_{m-1,l_k}^{-1}$ , we find that

$$\begin{aligned}
& \nabla \cdot (\hat{\zeta}_{m,l_k} \zeta_{m,k} \tilde{\psi}_{m,k} \sigma + \kappa_m \nabla \chi_{m,k} \circ X_{m-1,l_k}^{-1}) \\
&= (\nabla \cdot (\hat{\zeta}_{m,l_k} \zeta_{m,k} \psi_{m,k} \sigma + \kappa_m \nabla \chi_{m,k})) \circ X_{m-1,l_k}^{-1} \\
&\quad + \nabla \cdot ((\mathbf{I}_2 - \nabla X_{m-1,l_k} \circ X_{m-1,l_k}^{-1}) (\hat{\zeta}_{m,l_k} \zeta_{m,k} \psi_{m,k} \sigma + \kappa_m \nabla \chi_{m,k}) \circ X_{m-1,l_k}^{-1}). \quad (5.15)
\end{aligned}$$

Now comes the crucial point at which the corrector equation (3.13) for  $\chi_{m,k} = \chi_{m,k}^{\kappa_m}$  is used. By (3.13) and the fact that, similar to (5.13),

$$\nabla \cdot (\psi_{m,k} \sigma \nabla \chi_{m,k}^{\kappa_m}) = \nabla \psi_{m,k} \cdot \nabla^\perp \chi_{m,k}^{\kappa_m} = 0,$$

we can rewrite the first term of the right side of (5.15) as

$$(\nabla \cdot (\hat{\zeta}_{m,l_k} \zeta_{m,k} \psi_{m,k} \sigma + \kappa_m \nabla \chi_{m,k})) \circ X_{m-1,l_k}^{-1} = \partial_t \chi_{m,k} \circ X_{m-1,l_k}^{-1}.$$

Combining the previous displays, we obtain

$$\begin{aligned} & \nabla \cdot ((\kappa_m \mathbf{I}_2 + \hat{\zeta}_{m,l_k} \zeta_{m,k} \tilde{\psi}_{m,k} \sigma) (\mathbf{I}_2 + \nabla \tilde{\chi}_{m,k})) \\ &= \partial_t \chi_{m,k} \circ X_{m-1,l_k}^{-1} + \nabla \cdot (\kappa_m (\nabla X_{m-1,l_k}^{-1} - \mathbf{I}_2) \nabla \chi_{m,k} \circ X_{m-1,l_k}^{-1}) \\ & \quad + \nabla \cdot ((\mathbf{I}_2 - \nabla X_{m-1,l_k} \circ X_{m-1,l_k}^{-1}) (\hat{\zeta}_{m,l_k} \zeta_{m,k} \psi_{m,k} \sigma + \kappa_m \nabla \chi_{m,k}) \circ X_{m-1,l_k}^{-1}). \end{aligned} \quad (5.16)$$

Finally, using (5.16), we can compute the divergence of the last line of (5.12), which is also equal to negative of the first line of (5.11):

$$\begin{aligned} & \nabla \cdot \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} (\kappa_m \mathbf{I}_2 + \hat{\zeta}_{m,l_k} \zeta_{m,k} \tilde{\psi}_{m,k} \sigma) (\mathbf{I}_2 + \nabla \tilde{\chi}_{m,k}) \nabla (T_{m-1} \circ X_{m-1,k}) \circ X_{m-1,k}^{-1} \\ &= \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} \partial_t \chi_{m,k} \circ X_{m-1,k}^{-1} \cdot \nabla (T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1} \\ & \quad + \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} (\nabla \cdot (\kappa_m (\nabla X_{m-1,k}^{-1} - \mathbf{I}_2) \nabla \chi_{m,k} \circ X_{m-1,k}^{-1})) \cdot \nabla (T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1} \\ & \quad + \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} \nabla \cdot ((\mathbf{I}_2 - \nabla X_{m-1,l_k} \circ X_{m-1,l_k}^{-1}) (\hat{\zeta}_{m,l_k} \zeta_{m,k} \psi_{m,k} \sigma + \kappa_m \nabla \chi_{m,k}) \circ X_{m-1,l_k}^{-1}) \\ & \quad \cdot \nabla (T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1} \\ & \quad + \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} (\kappa_m \mathbf{I}_2 + \hat{\zeta}_{m,l_k} \zeta_{m,k} \tilde{\psi}_{m,k} \sigma) (\mathbf{I}_2 + \nabla \tilde{\chi}_{m,k}) : \sum_{l \in \mathbb{Z}} \hat{\xi}_{m,l} \nabla (\nabla (T_{m-1} \circ X_{m-1,l}) \circ X_{m-1,l}^{-1}). \end{aligned} \quad (5.17)$$

This expression will be substituted for the first line of (5.11), in view of (5.12), which will then be substituted for the second term of (5.2). The second and third terms on the right side of (5.17) are acceptable errors which are estimated in the next subsection. The first term on the right side of (5.17) will cancel the second term on the right side of (5.8) when we combine (5.8) and (5.11). The last term will be combined with the term involving  $\mathbf{J}_m$  on the last line of (5.8), which centers its mean and renders the resulting expression an acceptable error.

*5.1.3. Formula for the error of the ansatz.* We now combine (5.8) and (5.11), (5.12) and (5.17) to obtain an explicit expression for  $(\partial_t - \kappa_m \Delta + \mathbf{b}_m \cdot \nabla) \tilde{\theta}_m$ . As we merge (5.8) and (5.11), we recall that the second line of (5.8) cancels the first line of (5.17), and that the last line of (5.8) can be nicely combined with the last line of (5.17).

The result is the following equation satisfied by the ansatz  $\tilde{\theta}_m$ :

$$(\partial_t - \kappa_m \Delta + \mathbf{b}_m \cdot \nabla) \tilde{\theta}_m = \sum_{k \in 2\mathbb{Z}+1} (\partial_t \xi_{m,k} \tilde{\chi}_{m,k}) \cdot \nabla (T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1} \quad (5.18)$$

$$+ \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} \tilde{\chi}_{m,k} \cdot (\nabla X_{m-1,l_k} \circ X_{m-1,l_k}^{-1}) \nabla (\nabla \cdot (\mathbf{K}_m + \mathbf{s}_{m-1}) \nabla T_{m-1}) \quad (5.19)$$

$$- \nabla \cdot \sum_{l \in \mathbb{Z}} \hat{\xi}_{m,l} (\kappa_m \mathbf{I}_2 + \tilde{\psi}_m \sigma) (\nabla T_{m-1} - \nabla (T_{m-1} \circ X_{m-1,l}) \circ X_{m-1,l}^{-1}) \quad (5.20)$$

$$- \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} \nabla \cdot (\kappa_m (\nabla X_{m-1,l_k}^{-1} - \mathbf{I}_2) \nabla \chi_{m,k} \circ X_{m-1,l_k}^{-1}) \nabla (T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1} \quad (5.21)$$

$$- \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} \nabla \cdot \left( (\mathbf{I}_2 - \nabla X_{m-1,l_k} \circ X_{m-1,l_k}^{-1}) (\hat{\zeta}_{m,l_k} \zeta_{m,k} \tilde{\psi}_{m,k} \sigma + \kappa_m \nabla \chi_{m,k} \circ X_{m-1,l_k}^{-1}) \right) \times \nabla (T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1} \quad (5.22)$$

$$- \nabla \cdot \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} (\kappa_m \mathbf{I}_2 + \tilde{\psi}_m \sigma) \tilde{\chi}_{m,k} \nabla (\nabla (T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1}) \quad (5.23)$$

$$- \nabla \cdot (\kappa_m \mathbf{I}_2 + \tilde{\psi}_m \sigma) \nabla \tilde{H}_m \quad (5.24)$$

$$+ \nabla \cdot (\mathbf{d}_m + \mathbf{e}_{m-1}) + \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} \tilde{\chi}_{m,k} (\nabla X_{m-1,l_k} \circ X_{m-1,l_k}^{-1}) \cdot \nabla (\nabla \cdot \mathbf{e}_{m-1}) \quad (5.25)$$

$$- \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} \left( \mathbf{J}_m - (\kappa_m \mathbf{I}_2 + \hat{\zeta}_{m,l_k} \zeta_{m,k} \tilde{\psi}_{m,k} \sigma) (\mathbf{I}_2 + \nabla \tilde{\chi}_{m,k}) \right) : \sum_{l \in \mathbb{Z}} \hat{\xi}_{m,l} \nabla (\nabla (T_{m-1} \circ X_{m-1,l}) \circ X_{m-1,l}^{-1}) \quad (5.26)$$

The term (5.18) is due to the time cutoff  $\xi_{m,k}$  and is extremely small by (3.11), as we will see. The terms in (5.19)–(5.22) are errors caused by the “twisting” introduced into our ansatz by the composing  $T_{m-1}$  with the flows  $X_{m-1,l}$  and inverse flows  $X_{m-1,l}^{-1}$ ; these will be controlled using the bounds on the flows and inverse flows we proved in Section 2: see (2.40) and (2.67). The terms (5.23), (5.24) and (5.26) are “routine homogenization errors” which are expected: these are small because the correctors are small, as we find in error estimates in classical homogenization theory. Finally, the very tiny error terms in (5.25) is reflection of the way we constructed  $T_{m-1}$  and  $\tilde{H}_m$ , which actually solve the equations (4.13) and (4.22), respectively, which are slight perturbations of the equations (4.8) and (4.16) we initially wanted them to solve.

In the next subsection, we will show that the  $L_t^2 \dot{H}_x^{-1}$  norm of each of the numerous error terms on the right side of the expression above is at most  $O(\kappa_m^{1/2} \varepsilon_m^\delta)$ . This will subsequently permit us to deduce that the  $L_t^\infty L_x^2$  difference of the ansatz  $\tilde{\theta}_m$  and the true solution  $\theta_m$  of (4.3) is small, and thus to compare the decay in time of the  $L_x^2$  norm of  $\theta_m$  with that of  $\theta_{m-1}$ , up to a suitable error.

## 5.2 Estimates of the nine error terms

In this subsection, we prove the existence of a constant  $C(\beta) < \infty$  such that, if the parameter  $\Lambda$  satisfies

$$\Lambda \geq C \quad (5.27)$$

then, for every  $m \in \mathbb{N}$  satisfying

$$m \geq m_{\theta_0} := \inf \left\{ n \in \mathbb{N} : n \geq 2, \varepsilon_{n-1}^{1+\gamma/2} \leq R_{\theta_0} \right\}, \quad (5.28)$$

we have the estimate

$$\kappa_m^{-1/2} \|(\partial_t - \kappa_m \Delta + \mathbf{b}_m \cdot \nabla) \tilde{\theta}_m\|_{L^2((0,1); \dot{H}^{-1}(\mathbb{T}^2))} \leq C \varepsilon_{m-1}^\delta \|\theta_0\|_{L^2(\mathbb{T}^2)}. \quad (5.29)$$

The proof of (5.29) amounts to showing that each of the error terms in (5.18)–(5.26) can be estimated by the right side of (5.29).

**Remark 5.1.** The role of the restrictions  $\Lambda \geq C$  and  $m \leq m_{\theta_0}$  is to ensure that the conditions in (4.61) are valid and, therefore, all of the estimates on  $T_{m-1}$  and  $\tilde{H}_m$  proved in Section 4 are in force.

*5.2.1. Application of ergodic lemma for controlling nondivergence form terms.* For each of the four nondivergence terms in (5.19), (5.21), (5.22) and (5.26), we are faced with estimating the  $L_t^2 \dot{H}_x^{-1}$  norm of a function of the form  $f(g \circ Z^{-1})$ , which is the product of a “fast” function  $g \circ Z^{-1}$ , the composition of an  $\varepsilon_m$ -periodic, mean-zero function  $g$  with a smooth inverse flow  $Z^{-1}$  which has analyticity radius  $\varepsilon_{m-1}$ , and a “slow” function  $f$  which is 1-periodic, smooth and has analyticity radius of order  $\varepsilon_{m-1}$ . We expect that a weak norm of  $f(g \circ Z^{-1})$  would inherit smallness from the relatively fast oscillations of  $g$ , which modulate the slower signal  $f$ , up to an error which depends on the scale separation between  $\varepsilon_m$  and  $\varepsilon_{m-1}$ . Thanks to the periodicity of  $g$ , this error turns out to be exponentially small in the ratio  $\varepsilon_{m-1}/\varepsilon_m$ , as we show in Appendix C. Roughly, what we have is that

$$\begin{aligned} & \|f(g \circ Z^{-1})\|_{L^2((0,1); \dot{H}^{-1}(\mathbb{T}^2))} \\ & \leq C \varepsilon_m \|f\|_{L^2((0,1) \times \mathbb{T}^2)} \|g\|_{L^2((0,1) \times \mathbb{T}^2)} + C \|g\|_{L^2((0,1) \times \mathbb{T}^2)} \exp\left(-\frac{\varepsilon_{m-1}}{C \varepsilon_m}\right). \end{aligned} \quad (5.30)$$

The exact statement can be found in Remark C.5.

The first term in (5.30) represents the scaling of the  $\dot{H}^{-1}(\mathbb{T}^2)$  norm compared to the  $L^2(\mathbb{T}^2)$  norm for a mean-zero,  $\varepsilon_m$ -periodic function. For instance, in the case  $f \equiv 1$  and  $Z = \mathbf{I}_d$  and  $g$  is constant in time, we have  $\|g\|_{\dot{H}^{-1}(\mathbb{T}^2)} \simeq \varepsilon_m \|g\|_{L^2((0,1) \times \mathbb{T}^2)}$ . So essentially what (5.30) says in fact is that the modulation by  $f$  and  $Z$  does not alter this estimate, up to an error which is exponentially small in the scale separation. The constants  $C$  in the second term on the right of (5.30) depend on the appropriate analyticity norms of  $f$  and  $Z$ .

Here we check the applicability of this estimate to the terms (5.19), (5.21), (5.22) and (5.26). In all cases, the role of the measure-preserving mapping  $Z$  is played by  $X_{k-1,m}$  and it needs only to be estimated for times in the support of the cutoff function  $\xi_{m,k}$ . The required analyticity bounds are then a consequence of Corollary 2.4, in particular the estimate (2.70), which implies, in view of (2.12) and the definition of  $\xi_{m,k}$ ,

$$\sup_{t \in \text{supp } \xi_{m,k}} \max_{|\alpha|=n} \|\partial^\alpha X_{m-1,k}(t, \cdot)\|_{L^\infty(\mathbb{R}^2)} \leq C n! (C \varepsilon_{m-1}^{-1})^{n-1}, \quad \forall n \in \mathbb{N}. \quad (5.31)$$

The choices of  $f$  and  $g$  we need to make are different in each estimate, but in every situation we show that our choice of  $f$  satisfies, for some constant  $C < \infty$ ,

$$\sup_{t \in \text{supp } \xi_{m,k}} \max_{|\alpha|=n} \|\partial^\alpha f(t, \cdot)\|_{L^\infty(\mathbb{R}^2)} \leq C n! (C \varepsilon_{m-1}^{-1})^{n+1}, \quad \forall n \in \mathbb{N}. \quad (5.32)$$

We then may apply Remark C.5 with  $R = r = C_X = C\varepsilon_{m-1}^{-1}$  and  $N = \varepsilon_m^{-1}$ . The result (C.14) then yields

$$\begin{aligned} & \|f(g \circ Z^{-1})\|_{L^2((0,1);\dot{H}^{-1}(\mathbb{T}^2))} \\ & \leq C\varepsilon_m \|f\|_{L^2((0,1) \times \mathbb{T}^2)} \|g\|_{L^2((0,1) \times \mathbb{T}^2)} + C\|g\|_{L^2((0,1) \times \mathbb{T}^2)} \varepsilon_{m-1}^{-1} \exp\left(-\frac{\varepsilon_{m-1}}{C\varepsilon_m}\right) \\ & \leq C\varepsilon_m \|f\|_{L^2((0,1) \times \mathbb{T}^2)} \|g\|_{L^2((0,1) \times \mathbb{T}^2)} + C\varepsilon_{m-1}^{500} \|g\|_{L^2((0,1) \times \mathbb{T}^2)}. \end{aligned} \quad (5.33)$$

In the last line we used that, by (2.10),

$$\varepsilon_{m-1}^{-1} \exp\left(-\frac{\varepsilon_{m-1}}{C\varepsilon_m}\right) \leq \varepsilon_{m-1}^{-1} \exp(-c\varepsilon_{m-1}^{-(q-1)}) \leq C\varepsilon_{m-1}^{500}.$$

For (5.19), we use the choices

$$f = \xi_{m,k} (\nabla X_{m-1,l_k} \circ X_{m-1,l_k}^{-1}) \nabla \nabla \cdot ((\mathbf{K}_m + \mathbf{s}_{m-1}) \nabla T_{m-1}) \quad \text{and} \quad g = \chi_{m,k}. \quad (5.34)$$

The desired bound for  $f$  in (5.32) is a consequence of (2.68), the bounds for  $\mathbf{s}_{m-1}$  in (4.39), the estimate for  $\kappa_{m-1}$  in (3.45), Lemma 4.3 and the product estimate of Lemma B.1.

The term (5.21) is not in divergence form, nor is it entirely nondivergence form. Therefore, we split it into the sum of two terms, one in nondivergence form and the other in divergence form:

$$\begin{aligned} (5.21) &= \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} \kappa_m (\nabla X_{m-1,l_k}^{-1} - \mathbf{I}_2) \nabla \chi_{m,k} \circ X_{m-1,l_k}^{-1} \nabla (\nabla (T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1}) \\ &\quad - \nabla \cdot \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} \kappa_m (\nabla X_{m-1,l_k}^{-1} - \mathbf{I}_2) (\nabla \chi_{m,l_k} \circ X_{m-1,l_k}^{-1}) \nabla (T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1}. \end{aligned} \quad (5.35)$$

Only the first term in nondivergence form requires the use of the ergodic lemma. We apply it with the choices

$$f = \xi_{m,k} \kappa_m (\nabla X_{m-1,k}^{-1} - \mathbf{I}_2) \nabla (\nabla (T_{m-1} \circ X_{m-1,k}) \circ X_{m-1,k}^{-1}) \quad \text{and} \quad g = \nabla \chi_{m,k}. \quad (5.36)$$

Recall that

$$\nabla (T_{m-1} \circ X_{m-1,k}) \circ X_{m-1,k}^{-1} = (\nabla X_{m-1,k} \circ X_{m-1,k}^{-1}) \nabla T_{m-1}. \quad (5.37)$$

Therefore the desired bound for  $f$  in (5.32) is a consequence of (2.40), (2.68) and the product estimate of Lemma B.1.

We next consider the term (5.22) which, as for (5.21), must be split since it is not completely in nondivergence form. In order to write our expressions more compactly, we denote

$$Y_{m-1,l} := \mathbf{I}_2 - \nabla X_{m-1,l} \circ X_{m-1,l}^{-1} \quad (5.38)$$

and we write (5.22) as

$$\begin{aligned} (5.22) &= \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} Y_{m-1,l_k} (\zeta_{m,k} \tilde{\psi}_{m,k} \sigma + \kappa_m \nabla \chi_{m,k} \circ X_{m-1,l_k}^{-1}) \nabla (\nabla (T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1}) \\ &\quad - \nabla \cdot \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} Y_{m-1,k} (\zeta_{m,k} \tilde{\psi}_{m,k} \sigma + \kappa_m \nabla \chi_{m,k} \circ X_{m-1,l_k}^{-1}) \nabla (T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1}. \end{aligned} \quad (5.39)$$

Only the first term on the right of (5.39) will require the use of the ergodic lemma, and we will apply it with the choice

$$\begin{cases} f = \xi_{m,k} (I_2 - \nabla X_{m-1,l_k} \circ X_{m-1,l_k}^{-1}) \nabla (\nabla (T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1}), & \text{and} \\ g = \zeta_{m,k} \psi_{m,k} \sigma + \kappa_m \nabla \chi_{m,k}. \end{cases} \quad (5.40)$$

In view of (5.37), the desired analyticity estimate (5.32) for  $f$  is a consequence of estimate (2.68) of Corollary 2.4, Lemma 4.3 and the product estimate of Lemma B.1.

Finally, for (5.26), we apply the ergodic lemma (with the time variable frozen) with

$$\begin{cases} f = \xi_{m,k} \hat{\xi}_{m,l} \nabla (\nabla (T_{m-1} \circ X_{m-1,l}) \circ X_{m-1,l}^{-1}), \\ g = \mathbf{J}_m - (\kappa_m I_2 + \hat{\zeta}_{m,l_k} \zeta_{m,k} \psi_{m,k} \sigma) (I_2 + \nabla \chi_{m,k}). \end{cases} \quad (5.41)$$

In view of (5.37), the desired analyticity estimate (5.32) for  $f$  is a consequence of estimate (2.68) of Corollary 2.4, Lemma 4.3 and the product estimate of Lemma B.1.

*5.2.2. A reference list of basic estimates.* The following identities and estimates, proved above and collected here for the convenience of the reader, will be used repeatedly in the estimates of the terms on the right side of the Big Display on Page 73:

$$\varepsilon_m \simeq \varepsilon_{m-1}^q, \quad (5.42)$$

$$a_m = \varepsilon_m^{\beta-2}, \quad (5.43)$$

$$\tau_m \simeq a_{m-1}^{-1} \varepsilon_{m-1}^{2\delta}, \quad (5.44)$$

$$\kappa_m \simeq a_m \varepsilon_m^{2+\gamma} = \varepsilon_m^{\beta+\gamma} = \varepsilon_m^{\frac{2q}{q+1}\beta}, \quad (5.45)$$

$$\frac{\varepsilon_m^2}{\kappa_m \tau_m} \lesssim \varepsilon_{m-1}^{2\delta}, \quad (5.46)$$

$$\|\chi_{m,k}^\kappa\|_{L^\infty(\mathbb{R} \times \mathbb{T}^2)} + \varepsilon_m \|\nabla \chi_{m,k}^\kappa\|_{L^\infty(\mathbb{R} \times \mathbb{T}^2)} \lesssim \frac{a_m \varepsilon_m^3}{\kappa_m} \simeq \varepsilon_m^{1-\gamma}, \quad (5.47)$$

$$\|\tilde{\psi}_m\|_{L^\infty(\mathbb{R} \times \mathbb{T}^2)} \simeq \|\psi_{m,k}\|_{L^\infty(\mathbb{R} \times \mathbb{T}^2)} \simeq a_m \varepsilon_m^2 = \varepsilon_m^\beta, \quad (5.48)$$

$$\|\nabla \tilde{H}_m(t, \cdot)\|_{L^2([0,1] \times \mathbb{T}^2)} \lesssim \varepsilon_{m-1}^{4\delta} \kappa_{m-1}^{-1/2} \|\theta_0\|_{L^2(\mathbb{T}^2)}, \quad (5.49)$$

$$\|\nabla^{n+1} T_{m-1}\|_{L^2((0,1) \times \mathbb{T}^d)} \lesssim n! (C \varepsilon_{m-1}^{-(1+\gamma/2)})^n \kappa_{m-1}^{-1/2} \|\theta_0\|_{L^2(\mathbb{T}^2)}, \quad \forall n \in \mathbb{N}. \quad (5.50)$$

In each of the above inequalities, the symbols  $\lesssim$  and  $\simeq$  are to be interpreted as asserting inequalities and two-sided inequalities, respectively, with implicit prefactor constants which depend only on the parameter  $\beta$ . For (5.42), (5.43) and (5.44), see (2.10), (2.11) and (2.12), respectively; for (5.45) see (3.45); for (5.46), see (3.46); for (5.47), see (3.10); for (5.49), see (4.80); for (5.48), see (2.20); finally, for (5.50), see Lemma 4.3.

We present the estimates for the terms on the right side of the Big Display on Page 73, in consecutive order.

*The estimate of (5.18).* The smallness of this term is a reflection of the fact that the spatial oscillations in the shear flow homogenize on time scales of order  $\varepsilon_m^2/\kappa_m$ , which is much less than  $\tau_m$  by a factor of  $\varepsilon_{m-1}^{2\delta}$ . While it is in nondivergence form, we can bound it brutally using an exponential

factor so we do not need to use the ergodic lemma (Remark C.5) in the appendix; an  $L^2$  estimate is more than sufficient. The claim is that

$$\left\| \sum_{k \in 2\mathbb{Z}+1} (\partial_t \xi_{m,k} \tilde{\chi}_{m,k}) \cdot \nabla (T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1} \right\|_{L^2((0,1) \times \mathbb{T}^2)} \leq C \kappa_m^{1/2} \varepsilon_{m-1}^{500} \|\theta_0\|_{L^2(\mathbb{T}^2)}. \quad (5.51)$$

Observe that (3.11), (2.26) and (5.46) imply that  $\tilde{\chi}_{m,k}$  and  $\tilde{H}_m$  is very small on the support of  $\partial_t \xi_{m,k}$ . We can therefore estimate the left side of (5.51) by

$$\begin{aligned} \sup_{k \in 2\mathbb{Z}} \|\partial_t \xi_{m,k} \tilde{\chi}_{m,k}\|_{L^\infty(\mathbb{R} \times \mathbb{T}^d)} \|\nabla T_{m-1}\|_{L^2((0,1) \times \mathbb{T}^d)} &\leq C \exp(-c\varepsilon_{m-1}^{-2\delta}) \|\nabla \theta_{m-1}\|_{L^2((0,1) \times \mathbb{T}^d)} \\ &\leq C \kappa_m^{1/2} \underbrace{(\kappa_m^{-1/2} \kappa_{m-1}^{-1/2} \exp(-c\varepsilon_{m-1}^{-2\delta}))}_{\leq C\varepsilon_{m-1}^{500}} \|\theta_0\|_{L^2(\mathbb{T}^2)}. \end{aligned}$$

This is (5.51).  $\square$

*The estimates of (5.19).* We claim that

$$\begin{aligned} \left\| \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} \tilde{\chi}_{m,k} \cdot (\nabla X_{m-1,l_k} \circ X_{m-1,l_k}^{-1}) \nabla \nabla \cdot ((\kappa_{m-1} \mathbf{I}_2 + \mathbf{s}_{m-1}) \nabla T_{m-1}) \right\|_{L^2((0,1); \dot{H}^{-1}(\mathbb{T}^2))} \\ \leq C \kappa_m^{1/2} C \varepsilon_{m-1}^{8\delta} \|\theta_0\|_{L^2(\mathbb{T}^2)}. \end{aligned} \quad (5.52)$$

We will use the ergodic lemma of Remark C.5, as explained above. We apply (5.33) with  $f$  and  $g$  chosen as in (5.34) to find that the left side of (5.52) is bounded from above by the sum of

$$C \varepsilon_{m-1}^{500} \|g\|_{L^2((0,1) \times \mathbb{T}^2)} = C \varepsilon_{m-1}^{500} \|\chi_{m,k}\|_{L^2((0,1) \times \mathbb{T}^2)} \leq C \varepsilon_{m-1}^{500} \varepsilon_m^{1-\gamma} \leq C \varepsilon_{m-1}^{500},$$

and

$$\begin{aligned} C \varepsilon_m \|f\|_{L^2((0,1) \times \mathbb{T}^2)} \|g\|_{L^2((0,1) \times \mathbb{T}^2)} \\ \leq C \varepsilon_m \|\chi_{m,k}\|_{L^\infty} \sup_{k \in \mathbb{N}} \|\xi_{m,k} \nabla X_{m-1,l_k}\|_{L^\infty(\mathbb{R} \times \mathbb{T}^2)} \|\nabla \nabla \cdot ((\mathbf{K}_m + \mathbf{s}_{m-1}) \nabla T_{m-1})\|_{L^2((0,1) \times \mathbb{T}^2)} \\ \leq C \varepsilon_m \cdot \varepsilon_m^{1-\gamma} \cdot C \cdot \kappa_{m-1} \varepsilon_{m-1}^{-2-\gamma} \kappa_{m-1}^{-1/2} \|\theta_0\|_{L^2((0,1) \times \mathbb{T}^2)} \\ = C \kappa_m^{1/2} \underbrace{\left( \frac{\kappa_{m-1}}{\kappa_m} \right)^{1/2} \left( \frac{\varepsilon_m}{\varepsilon_{m-1}} \right)^2}_{\leq C\varepsilon_{m-1}^{8\delta}} \varepsilon_m^{-\gamma} \varepsilon_{m-1}^{-\gamma} \|\theta_0\|_{L^2(\mathbb{T}^2)} \leq C \kappa_m^{1/2} C \varepsilon_{m-1}^{8\delta} \|\theta_0\|_{L^2(\mathbb{T}^2)}. \end{aligned}$$

In the previous display, we used the size and regularity estimates for  $\mathbf{s}_{m-1}$  in (4.38) and (4.39) as well as (5.42), (5.45), (5.47), (2.67) and (5.50) and the fact that

$$\begin{aligned} \underbrace{\left( \frac{\kappa_{m-1}}{\kappa_m} \right)^{1/2}}_{\leq C\varepsilon_{m-1}^{-\gamma\beta}} \underbrace{\left( \frac{\varepsilon_m}{\varepsilon_{m-1}} \right)^2}_{\leq C\varepsilon_{m-1}^{2(q-1)}} \underbrace{\varepsilon_m^{-\gamma} \varepsilon_{m-1}^{-\gamma}}_{\leq C\varepsilon_{m-1}^{-(q-1)\beta}} &\leq C \varepsilon_{m-1}^{8\delta} \end{aligned}$$

since, by (2.5) and (2.7),

$$2(q-1) - \gamma\beta - (q-1)\beta = (q-1) \left( 2 - \frac{\beta}{q+1} (2q+1) \right) = 2(q-1) \left( 1 - \frac{2q+1}{2q+2} \beta \right) = 8\delta.$$

This completes the proof of (5.52).  $\square$

The estimate of (5.20). We will show that

$$\begin{aligned} \left\| \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} (\kappa_m \mathbf{I}_2 + \tilde{\psi}_m \sigma) (\nabla T_{m-1} - \nabla (T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1}) \right\|_{L^2((0,1) \times \mathbb{T}^2)} \\ \leq C \kappa_m^{1/2} \varepsilon_{m-1}^{2\delta} \|\theta_0\|_{L^2(\mathbb{T}^2)}. \end{aligned} \quad (5.53)$$

Using the formula

$$\nabla T_{m-1} - \nabla (T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1} = (\mathbf{I}_2 - \nabla X_{m-1,l_k} \circ X_{m-1,l_k}^{-1}) \nabla T_{m-1}$$

and the estimates (5.48), (2.68) and (4.9) with  $n = 0$ , we can bound the left side of (5.53) from above by

$$\begin{aligned} \|\kappa_m \mathbf{I}_2 + \tilde{\psi}_m \sigma\|_{L^\infty(\mathbb{R} \times \mathbb{T}^2)} \sup_{k \in \mathbb{N}} \|\xi_{m,k} (\mathbf{I}_2 - \nabla X_{m-1,l_k})\|_{L^\infty(\mathbb{R} \times \mathbb{T}^2)} \|\nabla T_{m-1}\|_{L^2((0,1) \times \mathbb{T}^2)} \\ \leq \underbrace{C(\kappa_m + a_m \varepsilon_m^2)}_{\leq C \kappa_m^{1/2} \kappa_{m-1}^{1/2}} \varepsilon_{m-1}^{2\delta} \|\nabla T_{m-1}\|_{L^2((0,1) \times \mathbb{T}^2)} \leq C \kappa_m^{1/2} \varepsilon_{m-1}^{2\delta} \|\theta_0\|_{L^2(\mathbb{T}^2)}. \end{aligned}$$

This yields (5.53).  $\square$

The estimate of (5.21). We use the identity (5.35) to split (5.21) into a divergence-form part and a nondivergence form part. The claimed estimates for these are as follows:

$$\begin{aligned} \left\| \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} \kappa_m (\nabla X_{m-1,l_k}^{-1} - \mathbf{I}_2) \nabla \chi_{m,k} \circ X_{m-1,l_k}^{-1} \nabla (T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1} \right\|_{L^2((0,1) \times \mathbb{T}^2)} \\ \leq C \kappa_m^{1/2} \varepsilon_{m-1}^{2\delta} \|\theta_0\|_{L^2(\mathbb{T}^2)} \end{aligned} \quad (5.54)$$

and

$$\begin{aligned} \left\| \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} \kappa_m (\nabla X_{m-1,l_k}^{-1} - \mathbf{I}_2) \nabla \chi_{m,k} \circ X_{m-1,l_k}^{-1} \nabla (\nabla (T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1}) \right\|_{L^2((0,1); \dot{H}^{-1}(\mathbb{T}^2))} \\ \leq C \kappa_m^{1/2} \varepsilon_{m-1}^{2\delta} \|\theta_0\|_{L^2(\mathbb{T}^2)}. \end{aligned} \quad (5.55)$$

For the  $L^2$  estimate (5.54), we use (2.40), (5.47) and (5.50) to bound the left side from above by

$$\begin{aligned} \sup_{k \in \mathbb{N}} \|\xi_{m,k} (\nabla X_{m-1,l_k}^{-1} - \mathbf{I}_2)\|_{L^\infty(\mathbb{R} \times \mathbb{T}^2)} \kappa_m \|\nabla \chi_{m,k}\|_{L^\infty(\mathbb{R} \times \mathbb{T}^2)} \|\nabla (T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1}\|_{L^2((0,1) \times \mathbb{T}^2)} \\ \leq C \varepsilon_{m-1}^{2\delta} \kappa_m \varepsilon_m^{-\gamma} \|\nabla T_{m-1}\|_{L^2((0,1) \times \mathbb{T}^2)} \\ = C \kappa_m^{1/2} \varepsilon_{m-1}^{2\delta} \underbrace{\left( \frac{\kappa_m}{\kappa_{m-1}} \right)^{1/2} \varepsilon_m^{-\gamma}}_{\leq C \text{ by (5.45) \& (2.7)}} \|\theta_0\|_{L^2(\mathbb{T}^2)} \leq C \kappa_m^{1/2} \varepsilon_{m-1}^{2\delta} \|\theta_0\|_{L^2(\mathbb{T}^2)}. \end{aligned} \quad (5.56)$$

This completes the proof of (5.54).

For (5.55), we use the ergodic lemma of Remark C.5, as explained above. We apply (5.33) with  $f$  and  $g$  chosen as in (5.36) to find that the left side of (5.55) is bounded from above by the sum of

$$C \varepsilon_{m-1}^{500} \|\nabla \chi_{m,k}\|_{L^2((0,1) \times \mathbb{T}^2)} \leq C \varepsilon_{m-1}^{500} \varepsilon_m^{-\gamma} \leq C \varepsilon_{m-1}^{499},$$

and

$$\begin{aligned}
& C\varepsilon_m \|\xi_{m,k} \kappa_m (\nabla X_{m-1,l_k}^{-1} - \mathbf{I}_2) \nabla (\nabla (T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1})\|_{L^2((0,1) \times \mathbb{T}^2)} \|\nabla \chi_{m,k}\|_{L^2((0,1) \times \mathbb{T}^2)} \\
& \leq C\varepsilon_m \kappa_m \sup_{k \in \mathbb{N}} \|\xi_{m,k} (\nabla X_{m-1,l_k}^{-1} - \mathbf{I}_2)\|_{L^\infty(\mathbb{R} \times \mathbb{T}^2)} \|\nabla \chi_{m,k}\|_{L^\infty(\mathbb{R} \times \mathbb{T}^2)} \\
& \quad \times \|\nabla (\nabla (T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1})\|_{L^2((0,1) \times \mathbb{T}^2)}.
\end{aligned}$$

The latter term is almost the same as in (5.56), the only differences being that we have one extra derivative on the  $T_m$  expression, which costs exactly  $\varepsilon_{m-1}^{-(1+\gamma/2)}$ , and we also have an extra factor of  $\varepsilon_m = \varepsilon_{m-1}^q$ . This more than compensates, since, in view of (2.7),

$$q - 1 > (q - 1) \frac{\beta}{2} > \gamma > \frac{1}{2}\gamma. \quad (5.57)$$

This proof of (5.55) is therefore complete.  $\square$

*The estimate of (5.22).* We will prove that

$$\|(\text{5.22})\|_{L^2((0,1); \dot{H}^{-1}(\mathbb{T}^2))} \leq C\kappa_m^{1/2} \varepsilon_{m-1}^{2\delta} \|\theta_0\|_{L^2(\mathbb{T}^2)}. \quad (5.58)$$

We use the identity (5.39) to split (5.22) into divergence and nondivergence form terms. The estimate for the first term on the right side of (5.39) is obtained with the help of the ergodic lemma in Appendix C, as explained below (5.39). We apply (5.33) with the choice of  $f$  and  $g$  given by (5.40). Up to the exponentially small error, the estimate is reduced therefore to

$$\begin{aligned}
& C\varepsilon_m \|f\|_{L^2((0,1) \times \mathbb{T}^2)} \|g\|_{L^2((0,1) \times \mathbb{T}^2)} \\
& \leq C\varepsilon_m \sup_{k \in \mathbb{N}} \|\xi_{m,k} (\nabla X_{m-1,l_k}^{-1} - \mathbf{I}_2)\|_{L^\infty(\mathbb{R} \times \mathbb{T}^2)} \|\nabla (\nabla (T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1})\|_{L^2((0,1) \times \mathbb{T}^2)} \\
& \quad \times \|\psi_m + \kappa_m |\nabla \chi_{m,k}|\|_{L^\infty(\mathbb{R} \times \mathbb{T}^2)} \\
& \leq C\kappa_m \varepsilon_m^{1-\gamma} \varepsilon_{m-1}^{-1-(1+\gamma/2)} \kappa_{m-1}^{-1/2} \|\theta_0\|_{L^2(\mathbb{T}^2)} \leq C\kappa_m^{1/2} \varepsilon_{m-1}^{q-1-\gamma/2+2\delta} \|\theta_0\|_{L^2(\mathbb{T}^2)}.
\end{aligned}$$

As above in (5.57), we have that  $q - 1 - \gamma/2 > 0$ , therefore the right side of the previous display is bounded from above by the right side of (5.58).

The second term on the right side of (5.39) is in divergence form, so we just need to estimate the  $L^2$  norm of what is under the divergence. The claim is that

$$\begin{aligned}
& \left\| \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} Y_{m-1,k} (\hat{\zeta}_{m,l_k} \zeta_{m,k} \tilde{\psi}_{m,k} \sigma + \kappa_m \nabla \chi_{m,k} \circ X_{m-1,l_k}^{-1}) \nabla (T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1} \right\|_{L^2((0,1) \times \mathbb{T}^2)} \\
& \leq C\kappa_m^{1/2} \varepsilon_{m-1}^{2\delta} \|\theta_0\|_{L^2(\mathbb{T}^2)}. \quad (5.59)
\end{aligned}$$

The proof of (5.59) is almost the same as (5.54). Compared to the latter, we use the bound

$$\|\tilde{\psi}_{m,k}\|_{L^\infty} + \kappa_m \|\nabla \chi_{m,k}\|_{L^\infty} \leq C\kappa_m \varepsilon_m^{-\gamma}$$

instead of just the bound for  $\kappa_m \|\nabla \chi_{m,k}\|_{L^\infty}$ , and we substitute the bound

$$\|\xi_{m,k} Y_{m-1,k}\|_{L^\infty(\mathbb{R} \times \mathbb{T}^2)} \leq C\varepsilon_{m-1}^{2\delta},$$

in place of the bound for  $\|\xi_{m,k} (\nabla X_{m-1,l_k}^{-1} - \mathbf{I}_2)\|_{L^\infty(\mathbb{R} \times \mathbb{T}^2)}$ , which is essentially the same. Recall that  $Y_{m-1,k}$  is defined in (5.38) and the above bound is a consequence of (2.67) and (5.44). The proof of (5.59) and hence of (5.58) is now complete.  $\square$

*The estimates of (5.23).* The claimed estimate is

$$\begin{aligned} \left\| \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} (\kappa_m \mathbf{I}_2 + \tilde{\psi}_m \sigma) \tilde{\chi}_{m,k} \nabla (\nabla (T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1}) \right\|_{L^2((0,1) \times \mathbb{T}^2)} \\ \leq C \kappa_m^{1/2} \varepsilon_{m-1}^{4\delta} \|\theta_0\|_{L^2(\mathbb{T}^2)}. \end{aligned} \quad (5.60)$$

Using (5.47), (5.48) and (5.50), we bound the left hand side of (5.60) by

$$\begin{aligned} & \|\kappa_m \mathbf{I}_2 + \psi_m \sigma\|_{L^\infty} \|\chi_{m,k}\|_{L^\infty} \sup_{k \in \mathbb{N}} \left\| \xi_{m,k} \nabla (\nabla (T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1}) \right\|_{L^2((0,1) \times \mathbb{T}^2)} \\ & \leq C \kappa_m \left( 1 + \frac{a_m \varepsilon_m^2}{\kappa_m} \right) \varepsilon_m^{1-\gamma} \varepsilon_{m-1}^{-(1+\gamma/2)} \kappa_{m-1}^{-1/2} \|\theta_0\|_{L^2(\mathbb{T}^2)} \\ & = C \kappa_m^{1/2} \underbrace{\kappa_m^{1/2} \left( 1 + \frac{a_m \varepsilon_m^2}{\kappa_m} \right)}_{\leq C \kappa_{m-1}^{1/2}} \underbrace{\left( \frac{\varepsilon_m}{\varepsilon_{m-1}} \right) \varepsilon_m^{-\gamma} \varepsilon_{m-1}^{-\gamma/2} \kappa_{m-1}^{-1/2} \|\theta_0\|_{L^2(\mathbb{T}^2)}}_{\leq C \varepsilon_{m-1}^{4\delta}} \leq C \kappa_m^{1/2} \varepsilon_{m-1}^{4\delta} \|\theta_0\|_{L^2(\mathbb{T}^2)}. \end{aligned}$$

In the above display, we used that

$$\left( \frac{\varepsilon_m}{\varepsilon_{m-1}} \right) \varepsilon_m^{-\gamma} \varepsilon_{m-1}^{-\gamma/2} \leq C \varepsilon_{m-1}^{q-1-q\gamma-\gamma/2} = C \varepsilon_{m-1}^{4\delta} \quad (5.61)$$

since, by (2.5) and (2.7),

$$q - 1 - q\gamma - \frac{1}{2}\gamma = (q-1) \left( 1 - \frac{\beta}{q+1} \left( q + \frac{1}{2} \right) \right) = (q-1) \left( 1 - \frac{2q+1}{2q+2} \beta \right) = 4\delta.$$

The proof of (5.60) is complete.  $\square$

*The estimate of (5.24).* The claimed estimate is

$$\|(\kappa_m \mathbf{I}_2 + \tilde{\psi}_m \sigma) \nabla \tilde{H}_m\|_{L^2((0,1) \times \mathbb{T}^2)} \leq C \kappa_m^{1/2} \varepsilon_{m-1}^\delta \|\theta_0\|_{L^2(\mathbb{T}^2)}. \quad (5.62)$$

This follows easily from (5.49), (5.48) and (5.50). Indeed, we have that

$$\begin{aligned} \|(\kappa_m \mathbf{I}_2 + \tilde{\psi}_m \sigma) \nabla \tilde{H}_m\|_{L^2((0,1) \times \mathbb{T}^2)} & \leq \|\kappa_m \mathbf{I}_2 + \psi_m \sigma\|_{L^\infty} \|\nabla \tilde{H}_m\|_{L^2((0,1) \times \mathbb{T}^2)} \\ & \leq C \kappa_m \underbrace{\left( 1 + \frac{a_m \varepsilon_m^2}{\kappa_m} \right)}_{\leq C \kappa_m^{1/2} \kappa_{m-1}^{1/2}} \varepsilon_{m-1}^{4\delta} \kappa_{m-1}^{-1/2} \|\theta_0\|_{L^2(\mathbb{T}^2)} \leq C \kappa_m^{1/2} \varepsilon_{m-1}^{4\delta} \|\theta_0\|_{L^2(\mathbb{T}^2)}. \end{aligned}$$

This completes the proof of (5.62).  $\square$

*The estimate of (5.25).* The claimed estimate is

$$\|(\text{5.25})\|_{L^2((0,1); \dot{H}^{-1}(\mathbb{T}^2))} \leq C \kappa_m^{1/2} \varepsilon_{m-1}^{2\delta} \|\theta_0\|_{L^2(\mathbb{T}^2)}. \quad (5.63)$$

We first show the following bound for  $\mathbf{d}_m$ :

$$\|\mathbf{d}_m\|_{L^2([0,1] \times \mathbb{T}^2)} \leq \kappa_m^{1/2} \varepsilon_{m-1}^{2\delta} \|\theta_0\|_{L^2(\mathbb{T}^2)}. \quad (5.64)$$

This bound follows from (4.88), upon taking  $N_*$  to be sufficiently large to ensure that

$$\kappa_{m-1}^{1/2} (C\varepsilon_{m-1}^\delta)^{N*/2} \leq \kappa_m^{1/2} \varepsilon_{m-1}^{2\delta}.$$

In turn, this estimate follows from  $C\varepsilon_{m-1}^{\delta/2} \leq 1$  and the definition of  $N_*$  in (2.6), which implies  $N_* \geq 8 + 4(q-1)(\beta+\gamma)\delta^{-1} = 8 + \frac{128q^2}{q-1}$ . The last equality follows from (2.3), (2.5), and (2.7).

We next prove that

$$\|\mathbf{e}_{m-1}\|_{L^2([0,1] \times \mathbb{T}^2)} \leq C\kappa_m^{1/2} \varepsilon_{m-1}^{2\delta} \|\theta_0\|_{L^2(\mathbb{T}^2)}. \quad (5.65)$$

The inequality (4.66) with  $n = 0$  says that

$$\|\mathbf{e}_{m-1}\|_{L^2((0,1) \times \mathbb{T}^2)} \leq CC_{N_*} \kappa_{m-1}^{1/2} (C\varepsilon_{m-1}^{2\delta})^{N*/2} \|\theta_0\|_{L^2(\mathbb{T}^2)}. \quad (5.66)$$

If  $N_*$  is sufficiently large to ensure that

$$\kappa_{m-1}^{1/2} (C\varepsilon_{m-1}^{2\delta})^{N*/2} \leq \kappa_m^{1/2} \varepsilon_{m-1}^{2\delta},$$

then (5.66) directly implies (5.65). Similar to the previous paragraph, the above estimate follows from the definition of  $N_*$  in (2.6), which implies that  $N_* \geq 8 + 128q^2(q-1)$ . Note that since  $N_*$  is chosen as in (2.6), the constant  $C_{N_*}$  also depends only on  $\beta$ , and this gives us (5.65).

The claimed estimate for the third and final term in (5.25) is

$$\left\| \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} \tilde{\chi}_{m,k} (\nabla X_{m-1,l_k} \circ X_{m-1,l_k}^{-1}) \cdot \nabla (\nabla \cdot \mathbf{e}_{m-1}) \right\|_{L^2((0,1) \times \mathbb{T}^2)} \leq C\varepsilon_{m-1}^{2\delta} \kappa_m^{1/2} \|\theta_0\|_{L^2(\mathbb{T}^2)}. \quad (5.67)$$

We proceed similarly as above. Using (2.67), (5.47) and (4.66) with  $n = 2$ , We have that

$$\begin{aligned} & \left\| \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} \tilde{\chi}_{m,k} (\nabla X_{m-1,l_k} \circ X_{m-1,l_k}^{-1}) \cdot \nabla (\nabla \cdot \mathbf{e}_{m-1}) \right\|_{L^2((0,1) \times \mathbb{T}^2)} \\ & \leq \sup_{k \in \mathbb{N}} \|\tilde{\chi}_{m,k}\|_{L^\infty(\mathbb{R} \times \mathbb{T}^2)} \|\xi_{m,k} (\nabla X_{m-1,l_k} \circ X_{m-1,l_k}^{-1})\|_{L^\infty(\mathbb{R} \times \mathbb{T}^2)} \|\nabla^2 \mathbf{e}_{m-1}\|_{L^2((0,1) \times \mathbb{T}^2)} \\ & \leq C\varepsilon_m^{1-\gamma} \cdot CC_{N_*} \kappa_{m-1}^{1/2} (C\varepsilon_{m-1}^{-2-\gamma}) (C\varepsilon_{m-1}^\delta)^{N_*} \|\theta_0\|_{L^2(\mathbb{T}^2)} \\ & \leq CC_{N_*} \left( \varepsilon_{m-1}^{-1-(q-1)(\beta-1)-2q\gamma} (C\varepsilon_{m-1}^\delta)^{N_*} \right) \kappa_m^{1/2} \|\theta_0\|_{L^2(\mathbb{T}^2)}. \end{aligned}$$

Arguing as above, we need to ensure that  $N_*$  was chosen so large that the term in parentheses is bounded by a constant times  $\varepsilon_{m-1}^{2\delta}$ . It suffices to take

$$N_* \geq 2 + \delta^{-1} (1 + (q-1)(\beta-1) - 2q\gamma),$$

which is clearly satisfied by the  $N_*$  defined in (2.6). This completes the proof of (5.67) and thus of (5.63).  $\square$

*The estimates of (5.26).* We will show that

$$\begin{aligned} & \left\| \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} \left( \mathbf{J}_m - (\kappa_m \mathbf{I}_2 + \hat{\zeta}_{m,l_k} \zeta_{m,k} \tilde{\psi}_{m,k} \sigma) (\mathbf{I}_2 + \nabla \tilde{\chi}_{m,k}) \right) \nabla (\nabla (T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1}) \right\|_{L_t^2 \dot{H}_x^{-1}} \\ & \leq C\kappa_m^{1/2} \varepsilon_{m-1}^{4\delta} \|\theta_0\|_{L^2(\mathbb{T}^2)}. \end{aligned} \quad (5.68)$$

This term is in nondivergence form, so we need to apply the ergodic lemma in Appendix C, as explained above. We apply (5.33) with the choice of  $f$  and  $g$  given by (5.41). Discarding the exponentially small error term, this reduces the estimate (5.68) to an estimate for

$$\begin{aligned}
& C\varepsilon_m \|f\|_{L^2((0,1)\times\mathbb{T}^2)} \|g\|_{L^2((0,1)\times\mathbb{T}^2)} \\
& \leq C\varepsilon_m \sup_{k\in\mathbb{N}} \|\xi_{m,k} \nabla (\nabla(T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1})\|_{L^2((0,1)\times\mathbb{T}^2)} \\
& \quad \times \left\| \mathbf{J}_m - (\kappa_m \mathbf{I}_2 + \hat{\zeta}_{m,l_k} \zeta_{m,k} \psi_{m,k} \sigma) (\mathbf{I}_2 + \nabla \chi_{m,k}) \right\|_{L^\infty(\mathbb{R}\times\mathbb{T}^2)} \\
& \leq C\varepsilon_m \cdot C\varepsilon_{m-1}^{-(1+\gamma/2)} \kappa_{m-1}^{-1/2} \|\theta_0\|_{L^2(\mathbb{T}^2)} \cdot C\kappa_m \varepsilon_m^{-2\gamma} \\
& \leq C\kappa_m^{1/2} \underbrace{\varepsilon_m^{1-\gamma} \varepsilon_{m-1}^{-(1+\gamma/2)}}_{\leq C\varepsilon_{m-1}^{4\delta}} \underbrace{\varepsilon_m^{-\gamma} \left( \frac{\kappa_m}{\kappa_{m-1}} \right)^{1/2}}_{\leq C} \|\theta_0\|_{L^2(\mathbb{T}^2)} \leq C\kappa_m^{1/2} \varepsilon_{m-1}^{4\delta} \|\theta_0\|_{L^2(\mathbb{T}^2)}.
\end{aligned}$$

Here we used (5.45), (5.47), (5.48) and (5.61) again. This completes the proof of (5.68).  $\square$

We have now estimated every one of the terms on the right side of the Big Display on Page 73 and shown that they are each bounded by the right side of (5.29). The proof of (5.29) is therefore complete.

### 5.3 Energy cascade down the scales

In this section, we complete the proof of Theorem 1.1. The main step remaining is to use the estimates from the previous section to obtain lower bounds on the energy dissipation of  $\theta_m$  in terms of that of  $\theta_{m-1}$ , thereby formalizing the expectation that energy is pushed by advection into smaller scales, down the inertial-convection subrange, until one finally reaches a scale small enough that molecular diffusivity dominates.

Recall that we have fixed a small parameter  $\kappa \in \mathcal{K}$ , with the set  $\mathcal{K}$  defined in (3.43), and chosen  $M \in \mathbb{N}$  satisfying (3.44); the finite sequence  $\kappa_M, \kappa_{M-1}, \dots, \kappa_0$  is defined by (3.42).

**Proposition 5.2** (Main induction step). *There exist constants  $C(\beta) < \infty$ , such that, if the minimal scale separation parameter  $\Lambda$  satisfies  $\Lambda \geq C$ , then the following statement is valid. For every  $R_{\theta_0} > 0$  and  $\theta_0 \in C^\infty(\mathbb{T}^2)$  with  $\langle \theta_0 \rangle = 0$  which satisfies the quantitative analyticity condition*

$$\max_{|\alpha|=n} \|\partial^\alpha \theta_0\|_{L^2(\mathbb{T}^2)} \leq \|\theta_0\|_{L^2(\mathbb{T}^2)} \frac{n!}{R_{\theta_0}^n}, \quad \forall n \in \mathbb{N}, \quad (5.69)$$

if we define

$$m_{\theta_0} := \min \left\{ m \in \mathbb{N} : m \geq 2, \varepsilon_{m-1}^{1+\gamma/2} \leq R_{\theta_0} \right\}, \quad (5.70)$$

then, for every  $m \in \{m_{\theta_0}, \dots, M\}$ , we have the estimates

$$\|\theta_m - \theta_{m-1}\|_{L^\infty((0,1); L^2(\mathbb{T}^2))} + \kappa_m^{1/2} \|\nabla \theta_m - \nabla \tilde{\theta}_m\|_{L^2((0,1)\times\mathbb{T}^2)} \leq C\varepsilon_{m-1}^\delta \|\theta_0\|_{L^2(\mathbb{T}^2)} \quad (5.71)$$

and

$$\left| \frac{\kappa_m \|\nabla \theta_m\|_{L^2((0,1)\times\mathbb{T}^2)}^2}{\kappa_{m-1} \|\nabla \theta_{m-1}\|_{L^2((0,1)\times\mathbb{T}^2)}^2} - 1 \right| \leq C\varepsilon_{m-1}^\delta. \quad (5.72)$$

The proof of Proposition 5.2 is based on (5.29), which implies the bound on  $\nabla\theta_m - \nabla\tilde{\theta}_m$  in (5.71). We then use this estimate to obtain the rest of the estimates in the proposition by doing the computations for  $\tilde{\theta}_m$  and then switching back to  $\theta_m$  with the triangle inequality.

We will assume that the constant  $C$  is large enough that the conditions (5.27) and (5.28) are valid. Therefore the estimate (5.29) proved in the previous subsection is also valid.

*The proof of the estimate of  $\theta_m - \tilde{\theta}_m$  in  $L_t^2 H_x^1$ .* We prove the estimate

$$\|\theta_m - \tilde{\theta}_m\|_{L^\infty((0,1);L^2(\mathbb{T}^2))} + \kappa_m^{1/2} \|\nabla\theta_m - \nabla\tilde{\theta}_m\|_{L^2((0,1)\times\mathbb{T}^2)} \leq C\varepsilon_{m-1}^\delta \|\theta_0\|_{L^2(\mathbb{T}^2)}. \quad (5.73)$$

Since  $\theta_m = \tilde{\theta}_m$  at  $t = 0$ , the energy estimate and (5.29) yield

$$\begin{aligned} & \|\theta_m - \tilde{\theta}_m\|_{L^\infty((0,1);L^2(\mathbb{T}^2))}^2 + \kappa_m \|\nabla\theta_m - \nabla\tilde{\theta}_m\|_{L^2((0,1)\times\mathbb{T}^2)}^2 \\ & \leq \frac{C}{\kappa_m} \|(\partial_t - \kappa_m \Delta + \mathbf{b}_m \cdot \nabla) \tilde{\theta}_m\|_{L^2((0,1);\dot{H}^{-1}(\mathbb{T}^2))}^2 \leq C\varepsilon_{m-1}^{2\delta} \|\theta_0\|_{L^2(\mathbb{T}^2)}. \end{aligned}$$

This is (5.73).  $\square$

*Estimate of the energy dissipation of  $\theta_m$  in terms of  $\theta_{m-1}$ .* We next prove that (5.72) holds. In view of (5.71), it is essentially equivalent to prove (5.72) with  $\tilde{\theta}_m$  in place of  $\theta_m$ . We therefore return to the identity (5.10) and rearrange it in the form

$$\begin{aligned} \nabla\tilde{\theta}_m &= \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} (\mathbf{I}_2 + \nabla\chi_{m,k}) \circ X_{m-1,l_k}^{-1} \nabla T_{m-1} \\ &+ \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} (\nabla X_{m-1,l_k}^{-1} - \mathbf{I}_2) (\nabla\chi_{m,k}) \circ X_{m-1,l_k}^{-1} \nabla T_{m-1} \\ &+ \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} \nabla\tilde{\chi}_{m,k} (\nabla X_{m-1,l_k} \circ X_{m-1,l_k}^{-1} - \mathbf{I}_2) \nabla T_{m-1} \\ &+ \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} \tilde{\chi}_{m,k} \nabla (\nabla (T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1}) + \nabla\tilde{H}_m. \end{aligned} \quad (5.74)$$

The first term on the right side of (5.74) makes the leading order contribution. We proceed by estimating the second through fifth terms on the right side of (5.74), showing that they are suitably small. For the second of these terms we use (2.67) and (5.44) in place of (4.65):

$$\begin{aligned} & \kappa_m^{1/2} \left\| \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} (\nabla X_{m-1,l_k}^{-1} - \mathbf{I}_2) (\nabla\chi_{m,k}) \circ X_{m-1,l_k}^{-1} \nabla T_{m-1} \right\|_{L^2((0,1)\times\mathbb{T}^2)} \\ & \leq C \kappa_m^{1/2} \sup_{k \in \mathbb{N}} \left\| \nabla X_{m-1,l_k}^{-1} - \mathbf{I}_2 \right\|_{L^\infty(\text{supp } \xi_{m,k} \times \mathbb{T}^2)} \|\xi_{m,k} \nabla\chi_{m,k}\|_{L^\infty(\mathbb{R} \times \mathbb{T}^2)} \|\nabla T_{m-1}\|_{L^2((0,1)\times\mathbb{T}^2)} \\ & \leq C \varepsilon_{m-1}^{2\delta} \underbrace{\left( \frac{\kappa_m}{\kappa_{m-1}} \right)^{1/2} \varepsilon_m^{-\gamma}}_{\leq C \text{ by (5.45) \& (5.47)}} \|\theta_0\|_{L^2(\mathbb{T}^2)} \leq C \varepsilon_{m-1}^{2\delta} \|\theta_0\|_{L^2(\mathbb{T}^2)}. \end{aligned}$$

For the third of these terms we use (2.67) and (5.44) in place of (4.65):

$$\begin{aligned}
& \kappa_m^{1/2} \left\| \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} \nabla \tilde{\chi}_{m,k} (\nabla X_{m-1,l_k} \circ X_{m-1,l_k}^{-1} - I_2) \nabla T_{m-1} \right\|_{L^2((0,1) \times \mathbb{T}^2)} \\
& \leq C \kappa_m^{1/2} \sup_{k \in \mathbb{N}} \|\xi_{m,k} \nabla \tilde{\chi}_{m,k}\|_{L^\infty(\mathbb{R} \times \mathbb{T}^2)} \|\nabla X_{m-1,l_k} - I_2\|_{L^\infty(\text{supp } \xi_{m,k} \times \mathbb{T}^2)} \|\nabla T_{m-1}\|_{L^2((0,1) \times \mathbb{T}^2)} \\
& \leq C \underbrace{\left( \frac{\kappa_m}{\kappa_{m-1}} \right)^{1/2} \varepsilon_m^{-\gamma} \varepsilon_{m-1}^{2\delta}}_{\leq C \text{ by (5.45) \& (5.47)}} \|\theta_0\|_{L^2(\mathbb{T}^2)} \leq C \varepsilon_{m-1}^{2\delta} \|\theta_0\|_{L^2(\mathbb{T}^2)}.
\end{aligned}$$

For the first term on the last line of (5.74), from (2.70), (5.47), and (5.50), we have

$$\begin{aligned}
& \kappa_m^{1/2} \left\| \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} \tilde{\chi}_{m,k} \nabla (\nabla (T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1}) \right\|_{L^2((0,1) \times \mathbb{T}^2)} \\
& \leq C \kappa_m^{1/2} \sup_{k \in \mathbb{N}} \|\xi_{m,k} \tilde{\chi}_{m,k}\|_{L^\infty(\text{supp } \xi_{m,k} \times \mathbb{T}^2)} \|\nabla^2 T_{m-1}\|_{L^2((0,1) \times \mathbb{T}^2)} \\
& \quad + C \kappa_m^{1/2} \sup_{k \in \mathbb{N}} \|\xi_{m,k} \tilde{\chi}_{m,k} \nabla^2 X_{m-1,l_k}\|_{L^\infty(\text{supp } \xi_{m,k} \times \mathbb{T}^2)} \|\nabla T_{m-1}\|_{L^2((0,1) \times \mathbb{T}^2)} \\
& \leq C \underbrace{\left( \frac{\kappa_m}{\kappa_{m-1}} \right)^{1/2} \varepsilon_m^{-\gamma}}_{\leq C \text{ by (5.45)}} \underbrace{\left( \frac{\varepsilon_m}{\varepsilon_{m-1}} \right) \varepsilon_{m-1}^{-\gamma/2} \|\theta_0\|_{L^2(\mathbb{T}^2)}}_{\leq C \varepsilon_{m-1}^{2(q-1)/3}} \leq C \varepsilon_{m-1}^{\frac{2}{3}(q-1)} \|\theta_0\|_{L^2(\mathbb{T}^2)}.
\end{aligned}$$

Finally, in view of the estimate for  $\nabla \tilde{H}_m$  in (4.80), we have

$$\kappa_m^{1/2} \|\nabla \tilde{H}_m(t, \cdot)\|_{L^2([0,1] \times \mathbb{T}^2)} \leq C \varepsilon_{m-1}^{\delta+\gamma} \|\theta_0\|_{L^2(\mathbb{T}^2)}. \quad (5.75)$$

By (5.74), the triangle inequality and the previous five displays, we therefore obtain that

$$\kappa_m^{1/2} \left\| \nabla \tilde{\theta}_m - \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} (I_2 + \nabla \chi_{m,k}) \circ X_{m-1,l_k}^{-1} \nabla T_{m-1} \right\|_{L^2((0,1) \times \mathbb{T}^2)} \leq C \varepsilon_{m-1}^{2\delta} \|\theta_0\|_{L^2(\mathbb{T}^2)}. \quad (5.76)$$

In view of (4.65), (5.71) and (5.76), in order to prove (5.72) we have left to show

$$\begin{aligned}
& \left| \kappa_m \left\| \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} (I_2 + \nabla \chi_{m,k}) \circ X_{m-1,l_k}^{-1} \nabla T_{m-1} \right\|_{L^2((0,1) \times \mathbb{T}^2)}^2 - \kappa_{m-1} \|\nabla T_{m-1}\|_{L^2((0,1) \times \mathbb{T}^2)}^2 \right| \\
& \leq C \varepsilon_{m-1}^{2\delta} \|\theta_0\|_{L^2(\mathbb{T}^2)}^2. \quad (5.77)
\end{aligned}$$

The proof of (5.77) is based on an application of an ergodic lemma in the appendix (Lemma C.3). Before we apply the lemma, we define the matrix

$$\mathbf{F}(t, x) := \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} (I_2 + \nabla \chi_{m,k}) \circ X_{m-1,l_k}^{-1}.$$

We can therefore break up the left side of (5.77) as follows:

$$\begin{aligned}
& \left| \kappa_m \left\| \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} (\mathbf{I}_2 + \nabla \chi_{m,k}) \circ X_{m-1,l_k}^{-1} \nabla T_{m-1} \right\|_{L^2((0,1) \times \mathbb{T}^2)}^2 - \kappa_{m-1} \|\nabla T_{m-1}\|_{L^2((0,1) \times \mathbb{T}^2)}^2 \right| \\
& \leq \kappa_m \left| \int_0^1 \|\mathbf{F}(t, \cdot) \nabla T_{m-1}(t, \cdot)\|_{L^2(\mathbb{T}^2)}^2 dt - \int_0^1 \int_{\mathbb{T}^2} \nabla T_{m-1}(t, x) \cdot \langle (\mathbf{F}^t \mathbf{F})(t, \cdot) \rangle \nabla T_{m-1}(t, x) dx dt \right| \\
& \quad + \int_0^1 \left| \kappa_m \langle (\mathbf{F}^t \mathbf{F})(t, \cdot) \rangle - \mathbf{J}_m(t) \right| \|\nabla T_{m-1}(t, \cdot)\|_{L^2(\mathbb{T}^2)}^2 dt \\
& \quad + \left| \int_0^1 \langle \nabla T_{m-1}(t, \cdot) \cdot \mathbf{J}_m(t) \nabla T_{m-1}(t, \cdot) \rangle dt - \int_0^1 \langle \nabla T_{m-1}(t, \cdot) \cdot \langle \langle \mathbf{J}_m \rangle \rangle \nabla T_{m-1}(t, \cdot) \rangle \right| \\
& \quad + \left| \int_0^1 \langle \nabla T_{m-1}(t, \cdot) \cdot \langle \langle \mathbf{J}_m \rangle \rangle \nabla T_{m-1}(t, \cdot) \rangle - \kappa_{m-1} \|\nabla T_{m-1}\|_{L^2((0,1) \times \mathbb{T}^2)}^2 \right| \tag{5.78}
\end{aligned}$$

The first and third terms on the right side of (5.78) are estimated using an ergodic lemma (Lemma C.3 in the appendix); for the first term, we apply the ergodic lemma in space only, and for the third term we apply the ergodic lemma in time only. The second term on the right side of (5.78) is a consequence of Lemma 3.2. The fourth term is actually equal to zero, because  $\kappa_m \mathbf{I}_2$  is equal to  $\langle \mathbf{J}_m \rangle$  by definition: see (3.16) and (3.42).

We next bound the second term on the right side of (5.78). We observe that

$$\mathbf{F}^t \mathbf{F} = \mathbf{I}_2 + \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} ((\nabla \chi_{m,k})^t + \nabla \chi_{m,k}) \circ X_{m-1,l_k}^{-1} + \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k}^2 ((\nabla \chi_{m,k})^t \nabla \chi_{m,k}) \circ X_{m-1,l_k}^{-1}.$$

Indeed, this follows from the fact that the supports of  $\xi_{m,k}$  and  $\xi_{m,k'}$  have nonempty intersection only if  $k, k' \in 2\mathbb{Z}+1$  satisfy  $|k - k'| \leq 2$  and, by (3.34), we have

$$|k - k'| = 2 \implies (\nabla \chi_{m,k} \circ X_{m-1,l_k}^{-1})^\top (\nabla \chi_{m,k'} \circ X_{m-1,l_{k'}}^{-1}) = 0. \tag{5.79}$$

Taking the average in space and multiplying by  $\kappa_m$  yields, in view of (3.35),

$$\begin{aligned}
\kappa_m \langle \mathbf{F}^t \mathbf{F} \rangle(t) &= \kappa_m \mathbf{I}_2 + \kappa_m \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k}^2(t) \langle ((\nabla \chi_{m,k})^t \nabla \chi_{m,k}) \circ X_{m-1,l_k}^{-1} \rangle(t) \\
&= \kappa_m \mathbf{I}_2 + \kappa_m \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k}^2(t) \langle (\nabla \chi_{m,k})^t \nabla \chi_{m,k} \rangle(t) = \mathbf{E}_m(t). \tag{5.80}
\end{aligned}$$

Therefore, thanks to Lemma 3.2, we obtain

$$\begin{aligned}
& \int_0^1 \left| \kappa_m \langle (\mathbf{F}^t \mathbf{F})(t, \cdot) \rangle - \mathbf{J}_m(t) \right| \|\nabla T_{m-1}(t, \cdot)\|_{L^2(\mathbb{T}^2)}^2 dt \\
& = \int_0^1 \left| \mathbf{E}_m(t, \cdot) - \mathbf{J}_m(t) \right| \|\nabla T_{m-1}(t, \cdot)\|_{L^2(\mathbb{T}^2)}^2 dt \\
& \leq \kappa_{m-1} \left( \frac{\varepsilon_m^2}{\kappa_m \tau_m} \right) \|\nabla T_{m-1}\|_{L^2((0,1) \times \mathbb{T}^2)}^2 \leq \varepsilon_{m-1}^{2\delta} \|\theta_0\|_{L^2(\mathbb{T}^2)}^2, \tag{5.81}
\end{aligned}$$

where we used (5.46) and (5.50) to get the last inequality.

We next estimate the first term on the right side of (5.78). The claimed estimate is

$$\begin{aligned}
& \kappa_m \left| \int_0^1 \|\mathbf{F}(t, \cdot) \nabla T_{m-1}(t, \cdot)\|_{L^2(\mathbb{T}^2)}^2 dt - \int_0^1 \int_{\mathbb{T}^2} \nabla T_{m-1}(t, x) \cdot \langle (\mathbf{F}^t \mathbf{F})(t, \cdot) \rangle \nabla T_{m-1}(t, x) dx dt \right| \\
& \leq \varepsilon_{m-1}^{500} \|\theta_0\|_{L^2(\mathbb{T}^2)}^2. \tag{5.82}
\end{aligned}$$

For this we apply an ergodic lemma from Appendix C, namely Lemma C.3 with the time variable frozen. We therefore fix  $t \in (0, 1)$ ,  $k \in 2\mathbb{Z} + 1$  and  $i, j \in \{1, 2\}$  and we use the following choices for the  $X$ ,  $f$  and  $g$  appearing in that lemma:

$$\begin{cases} X = X_{m-1, l_k}, \\ f = \partial_{x_i} T_{m-1}(t, \cdot) \partial_{x_j} T_{m-1}(t, \cdot), \\ g = \left( I_2 + \xi_{m,k} \left( (\nabla \chi_{m,k})^t + \nabla \chi_{m,k} \right) + \xi_{m,k}^2 \left( (\nabla \chi_{m,k})^t \nabla \chi_{m,k} \right) \right)_{ij}. \end{cases}$$

and with the following choices of the constants appearing the hypotheses of that lemma:

$$\begin{cases} C_X = C\varepsilon_{m-1}, \\ R = C\varepsilon_{m-1}^{-1}, \\ C_f = C_{\theta_0} R_{\theta_0}^{-1} \kappa_{m-1}^{-1/2} \\ r = R_{\theta_0}^{-1} \varepsilon_{m-1}^{1+\gamma/2} \\ N = \varepsilon_m^{-1}. \end{cases}$$

Let's check the hypotheses of Lemma C.3 in this situation. We first note that (C.6) is valid since (2.40). The hypothesis (C.7) is valid due to (4.62). The periodicity assumption for  $g$  is clear from the construction of the correctors  $\chi_{m,k}$ , as these are  $\varepsilon_m$ -periodic. Finally, the condition (C.8) is valid since

$$\frac{Nr}{R(r + dC_X)} = \frac{\varepsilon_{m-1}^{2+\gamma/2}}{C\varepsilon_m R_{\theta_0} (R_{\theta_0}^{-1} \varepsilon_{m-1}^{1+\gamma/2} + \varepsilon_{m-1})} \geq \frac{\varepsilon_{m-1}^{1+\gamma/2}}{CR_{\theta_0} \varepsilon_m} = \frac{\varepsilon_{m-1}^{-(q-1)(1-\frac{\beta}{2(q+1)})}}{CR_{\theta_0}} \geq \frac{\varepsilon_{m-1}^{-2(q-1)/3}}{CR_{\theta_0}}.$$

Clearly the expression on the right side is at least 1, provided that  $\Lambda$  is chosen sufficiently large. Note that, after summing over  $k \in 2\mathbb{Z} + 1$ , the function  $g \circ X^{-1}$  is equal to the  $ij$ th component of  $\mathbf{F}^t \mathbf{F}(t, \cdot)$ . Therefore, an application of the lemma, namely (C.9), gives us the bound

$$\begin{aligned} \kappa_m \left| \left\| \mathbf{F}(t, \cdot) \nabla T_{m-1}(t, \cdot) \right\|_{L^2(\mathbb{T}^2)}^2 - \int_{\mathbb{T}^2} \nabla T_{m-1}(t, x) \cdot \langle (\mathbf{F}^t \mathbf{F})(t, \cdot) \rangle \nabla T_{m-1}(t, x) dx \right| \\ \leq C \kappa_{m-1}^{1/2} \exp \left( -\frac{\varepsilon_{m-1}^{-2(q-1)/3}}{CR_{\theta_0}} \right) \|\theta_0\|_{L^2(\mathbb{T}^2)}^2 \leq \varepsilon_{m-1}^{500} \|\theta_0\|_{L^2(\mathbb{T}^2)}^2. \end{aligned}$$

Integrating over  $t$  yields (5.82).

Returning to (5.78), we consider the third term on the right side. Our claim is that

$$\begin{aligned} \left| \int_0^1 \left\langle \nabla T_{m-1}(t, \cdot) \cdot \mathbf{J}_m(t) \nabla T_{m-1}(t, \cdot) \right\rangle dt - \int_0^1 \left\langle \nabla T_{m-1}(t, \cdot) \cdot \langle \langle \mathbf{J}_m \rangle \rangle \nabla T_{m-1}(t, \cdot) \right\rangle dt \right| \\ \leq C \varepsilon_{m-1}^{2\delta} \|\theta_0\|_{L^2(\mathbb{T}^2)}^2. \quad (5.83) \end{aligned}$$

To see this, we first recall that the bound for  $\mathbf{J}_m - \hat{\mathbf{J}}_m$  in (3.25), together with (3.46), implies that

the left side of (5.83) is bounded from above as

$$\begin{aligned}
& \left| \int_0^1 \left\langle \nabla T_{m-1}(t, \cdot) \cdot \mathbf{J}_m(t) \nabla T_{m-1}(t, \cdot) \right\rangle dt - \int_0^1 \left\langle \nabla T_{m-1}(t, \cdot) \cdot \langle\langle \mathbf{J}_m \rangle\rangle \nabla T_{m-1}(t, \cdot) \right\rangle dt \right| \\
& \leq \frac{C a_m^2 \varepsilon_m^4}{\kappa_m} \left( \frac{\varepsilon_m^2}{\kappa_m \tau_m} \right)^{N_*} \|\nabla T_{m-1}\|_{L^2([0,1] \times \mathbb{T}^2)}^2 + \left| \int_0^1 (\hat{\mathbf{J}}_m(t) - \langle\langle \hat{\mathbf{J}}_m \rangle\rangle) : \left\langle \nabla T_{m-1}(t, \cdot) \otimes \nabla T_{m-1}(t, \cdot) \right\rangle dt \right| \\
& \leq C (C \varepsilon_{m-1}^{2\delta})^{N_*} \kappa_{m-1} \|\nabla T_{m-1}\|_{L^2([0,1] \times \mathbb{T}^2)}^2 + \left| \int_0^1 (\hat{\mathbf{J}}_m(t) - \langle\langle \hat{\mathbf{J}}_m \rangle\rangle) : \left\langle \nabla T_{m-1}(t, \cdot) \otimes \nabla T_{m-1}(t, \cdot) \right\rangle dt \right| \\
& \leq C (C \varepsilon_{m-1}^{2\delta})^{N_*} \|\theta_0\|_{L^2(\mathbb{T}^2)}^2 + \left| \int_0^1 (\hat{\mathbf{J}}_m(t) - \langle\langle \hat{\mathbf{J}}_m \rangle\rangle) : \left\langle \nabla T_{m-1}(t, \cdot) \otimes \nabla T_{m-1}(t, \cdot) \right\rangle dt \right|. \tag{5.84}
\end{aligned}$$

Next, we recall from (3.20) that  $\hat{\mathbf{J}}_m(t) - \langle\langle \hat{\mathbf{J}}_m \rangle\rangle$  is a zero-mean  $\tau_m''$ -periodic function of time (recall (2.15)), and so we may write

$$\hat{\mathbf{J}}_m(t) - \langle\langle \hat{\mathbf{J}}_m \rangle\rangle = \sum_{n=0}^{N_*-1} (L_{m,n}(t) \mathbf{j}_{m,n}(t) - \langle\langle L_{m,n} \mathbf{j}_{m,n} \rangle\rangle) = \partial_t \hat{\mathbf{Q}}_m(t)$$

where

$$\|\hat{\mathbf{Q}}_m\|_{L^\infty([0,1])} \leq C \kappa_{m-1} \tau_m'', \quad \text{and} \quad \hat{\mathbf{Q}}_m(0) = \hat{\mathbf{Q}}_m(1) = 0.$$

In light of the above two displays, we integrate by parts in time, use the identity

$$\begin{aligned}
& \int_0^1 \hat{\mathbf{Q}}_m(t) \partial_t \left\langle \nabla T_{m-1}(t, \cdot) \otimes \nabla T_{m-1}(t, \cdot) \right\rangle dt \\
& = \int_0^1 \hat{\mathbf{Q}}_m(t) \int_{\mathbb{T}^2} \left( \mathcal{D}_{t,m-1} \nabla T_{m-1} \otimes \nabla T_{m-1} + \nabla T_{m-1} \otimes \mathcal{D}_{t,m-1} \nabla T_{m-1} \right)(t, x) dx dt,
\end{aligned}$$

which follows since  $\nabla \cdot \mathbf{b}_{m-1} = 0$ , and appeal to the estimates (4.62) and (4.67) to deduce

$$\begin{aligned}
& \left| \int_0^1 (\hat{\mathbf{J}}_m(t) - \langle\langle \hat{\mathbf{J}}_m \rangle\rangle) : \left\langle \nabla T_{m-1}(t, \cdot) \otimes \nabla T_{m-1}(t, \cdot) \right\rangle dt \right| \\
& \leq C \|\hat{\mathbf{Q}}_m\|_{L^\infty([0,1])} \|\nabla T_{m-1}\|_{L^2([0,1] \times \mathbb{T}^2)} \|\mathcal{D}_{t,m-1} \nabla T_{m-1}\|_{L^2([0,1] \times \mathbb{T}^2)} \\
& \leq C \kappa_{m-1} \tau_m'' \cdot \kappa_{m-1}^{-1/2} \|\theta_0\|_{L^2(\mathbb{T}^2)} \cdot C \varepsilon_{m-1}^{3\delta} (\tau_m')^{-1} \kappa_{m-1}^{-1/2} \|\theta_0\|_{L^2(\mathbb{T}^2)} \leq C \varepsilon_{m-1}^{2\delta} \|\theta_0\|_{L^2(\mathbb{T}^2)}^2.
\end{aligned}$$

Combining this display with (5.84) and using that

$$C (C \varepsilon_{m-1}^{2\delta})^{N_*} \kappa_{m-1} \|\nabla T_{m-1}\|_{L^2([0,1] \times \mathbb{T}^2)}^2 \leq \varepsilon_{m-1}^{500} \|\theta_0\|_{L^2(\mathbb{T}^2)}^2$$

for  $N_*$  sufficiently large, we obtain (5.83).

Finally, we collect the bounds we have obtained for the terms on the right side of (5.78), namely (5.81), (5.82), (5.83), and obtain that

$$\begin{aligned}
& \left| \kappa_m \left\| \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} (\mathbf{I}_2 + \nabla \chi_{m,k}) \circ X_{m-1, l_k}^{-1} \nabla T_{m-1} \right\|_{L^2((0,1) \times \mathbb{T}^2)}^2 - \kappa_{m-1} \|\nabla T_{m-1}\|_{L^2((0,1) \times \mathbb{T}^2)}^2 \right| \\
& \leq C \varepsilon_{m-1}^{2\delta} \|\theta_0\|_{L^2(\mathbb{T}^2)}^2.
\end{aligned}$$

This concludes the proof of (5.77).  $\square$

The estimate of  $\tilde{\theta}_m - T_{m-1}$  in  $L_t^\infty L_x^2$ . We next prove the estimate

$$\|\tilde{\theta}_m - T_{m-1}\|_{L^\infty((0,1);L^2(\mathbb{T}^2))} \leq C\varepsilon_{m-1}^\delta \|\theta_0\|_{L^2(\mathbb{T}^2)}. \quad (5.85)$$

We proceed by bounding the last two terms on the right side of (4.24). First, using (2.26), (5.47), the fact that  $X_{m-1,l_k}^{-1}$  is volume preserving, and that  $|\nabla X_{m-1,l_k}| \leq 2$  on  $\text{supp}(\xi_{m,k})$ , we obtain

$$\begin{aligned} & \left\| \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} \tilde{\chi}_{m,k} (\nabla(T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1}) \right\|_{L^\infty((0,1);L^2(\mathbb{T}^2))} \\ & \leq \sup_{k \in \mathbb{N}} \|\tilde{\chi}_{m,k}\|_{L^\infty(\mathbb{R} \times \mathbb{T}^2)} \left\| \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} (\nabla(T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1}) \right\|_{L^\infty((0,1);L^2(\mathbb{T}^2))} \\ & \leq C\varepsilon_m^{1-\gamma} \sup_{k \in \mathbb{N}} \|\xi_{m,k}(\nabla T_{m-1}) \circ X_{m-1,l_k}\|_{L^\infty((0,1);L^2(\mathbb{T}^2))}. \end{aligned} \quad (5.86)$$

The right side of the above display requires that we bound  $\|(\nabla T_{m-1}) \circ X_{m-1,l_k}(t, \cdot)\|_{L^2(\mathbb{T}^2)}$  in  $L^\infty(0, 1)$ , instead of  $L^2(0, 1)$  (a bound that would have been available from (5.50)). Instead, we appeal to the fundamental theorem of calculus in time which in light of  $\xi_{m,k}((k - \frac{5}{4})\tau_m) = 0$ , gives

$$\begin{aligned} & \sup_{t \in \mathbb{R}} \|\xi_{m,k}(t)(\nabla T_{m-1}) \circ X_{m-1,l_k}(t, \cdot)\|_{L^2(\mathbb{T}^2)} \\ & \leq \int_{(k - \frac{5}{4})\tau_m}^{(k + \frac{5}{4})\tau_m} \left( \|\partial_t \xi_{m,k}\|_{L^\infty(\mathbb{R})} \|\nabla T_{m-1}(t, \cdot)\|_{L^2(\mathbb{T}^2)} + \|\xi_{m,k}\|_{L^\infty(\mathbb{R})} \|\mathcal{D}_{t,m-1} \nabla T_{m-1}(t, \cdot)\|_{L^2(\mathbb{T}^2)} \right) dt. \end{aligned}$$

By further appealing to (4.62) with  $n = 0$  and to (4.67) with  $n = 0$  and  $\ell = 1$ , and using also (2.27), we deduce from the previous display that

$$\begin{aligned} & \sup_{t \in \mathbb{R}} \|\xi_{m,k}(t)(\nabla T_{m-1}) \circ X_{m-1,l_k}(t, \cdot)\|_{L^2(\mathbb{T}^2)} \\ & \leq C\tau_m^{-1/2} \|\nabla T_{m-1}\|_{L^2(0,1;L^2(\mathbb{T}^2))} + C\tau_m^{1/2} \|\mathcal{D}_{t,m-1} \nabla T_{m-1}\|_{L^2(0,1;L^2(\mathbb{T}^2))} \\ & \leq C \left( \tau_m^{-1/2} + \tau_m^{1/2} (\tau'_m)^{-1} \right) \kappa_{m-1}^{-1/2} \|\theta_0\|_{L^2(\mathbb{T}^2)} \leq C\tau_m^{-1/2} \kappa_{m-1}^{-1/2} \|\theta_0\|_{L^2(\mathbb{T}^2)}. \end{aligned} \quad (5.87)$$

By combining (5.86) and (5.87), we deduce

$$\begin{aligned} & \left\| \sum_{k \in 2\mathbb{Z}+1} \xi_{m,k} \tilde{\chi}_{m,k} (\nabla(T_{m-1} \circ X_{m-1,l_k}) \circ X_{m-1,l_k}^{-1}) \right\|_{L^\infty((0,1);L^2(\mathbb{T}^2))} \\ & \leq C \underbrace{\varepsilon_m^{1-\gamma} \tau_m^{-1/2} \kappa_{m-1}^{-1/2}}_{\leq C\varepsilon_{m-1}^\delta} \leq C\varepsilon_{m-1}^\delta \|\theta_0\|_{L^2(\mathbb{T}^2)}. \end{aligned} \quad (5.88)$$

In the last inequality we used that

$$\varepsilon_m^{1-\gamma} \tau_m^{-1/2} \kappa_{m-1}^{-1/2} = \varepsilon_{m-1}^{q(1-\gamma)-1+\beta/2-2\delta} \varepsilon_{m-1}^{-\frac{q\beta}{q+1}} = \varepsilon_{m-1}^{q(1-\gamma)-1+\beta/2-\frac{q\beta}{q+1}-3\delta} \varepsilon_{m-1}^\delta$$

and, since  $\beta \in (1, 4/3)$ ,

$$q(1-\gamma) - 1 + \frac{\beta}{2} - \frac{q\beta}{q+1} - 3\delta = (q-1) + \beta \left( \frac{1}{2} - \frac{q^2}{q+1} \right) - \frac{3(q-1)^2}{4(q+1)(4q-1)} = \frac{(4-3\beta)^2}{80\beta-64} \geq 0.$$

Finally, by the estimate for  $H_m$  in (4.79), we have

$$\|\tilde{H}_m(t, \cdot)\|_{L_t^\infty L_x^2([0,1] \times \mathbb{T}^2)} \leq C\varepsilon_{m-1}^\delta \|\theta_0\|_{L^2(\mathbb{T}^2)}. \quad (5.89)$$

The previous two displays, (4.24) and the triangle inequality yield (5.85).  $\square$

The proof of Proposition 5.2 is now complete, as the estimate for the first term on the left side of (5.71) follows from (4.65), (5.73) and (5.85).

**Remark 5.3** (Uniform continuity in time of the scalar variance). By a simple interpolation argument, the estimate for the first term in (5.71) can be improved to obtain some positive regularity in the time variable. Indeed, we have that there exists an exponent  $\mu(\beta) > 0$  and a constant  $C(\beta) < \infty$  such that

$$\|\theta_m - \theta_{m-1}\|_{C^{0,\mu}([0,1]; L^2(\mathbb{T}^2))} \leq C\varepsilon_{m-1}^{\delta/2} \|\theta_0\|_{H^1(\mathbb{T}^2)}. \quad (5.90)$$

To prove (5.90), we first interpolate  $C^{0,\mu}$  between  $L^\infty$  and  $C^{0,1}$  and apply (5.71) to get

$$\begin{aligned} \|\theta_m - \theta_{m-1}\|_{C^{0,\mu}([t_0,1]; L^2(\mathbb{T}^2))} &\leq \|\theta_m - \theta_{m-1}\|_{L^\infty([0,1]; L^2(\mathbb{T}^2))}^{1-\mu} \|\partial_t(\theta_m - \theta_{m-1})\|_{L^\infty([t_0,1]; L^2(\mathbb{T}^2))}^\mu \\ &\leq C\varepsilon_{m-1}^{(1-\mu)\delta} \|\theta_0\|_{L^2(\mathbb{T}^2)}^{1-\mu} \|\partial_t(\theta_m - \theta_{m-1})\|_{L^\infty([t_0,1]; L^2(\mathbb{T}^2))}^\mu. \end{aligned} \quad (5.91)$$

The second factor may be estimated very crudely using standard parabolic Schauder estimates. These imply that, any equation of the form (1.1) with  $\mathbf{b}(t, \cdot)$  divergence-free and satisfying (1.3), the solution  $\theta^\kappa$  of (1.1) satisfies

$$\|\theta^\kappa\|_{C_t^{1+\alpha/2} C_x^{2+\alpha}((t_0,1) \times \mathbb{T}^d)} \leq C(t_0^2 \wedge \kappa)^{-\frac{d+2}{2\alpha}} \|\theta_0\|_{L^2(\mathbb{T}^2)}. \quad (5.92)$$

To get this, cover the torus  $\mathbb{T}^d$  with balls of radius  $(t_0^2 \wedge \kappa)^{-1/\alpha}$  and apply the interior Schauder estimates (after rescaling) in the corresponding parabolic cylinders. Giving up volume factors then leads to (5.92). (The power of  $\kappa$  which appears is not important, we just need that it is some negative power of  $\kappa$ .) Applying this to  $\theta_m$  and  $\theta_{m-1}$  separately and then inserting the bound into (5.91), we obtain

$$\|\theta_m - \theta_{m-1}\|_{C^{0,\mu}([t_0,1]; L^2(\mathbb{T}^2))} \leq C\varepsilon_{m-1}^{(1-\mu)\delta} (t_0^2 \wedge \varepsilon_m^{\beta+\gamma})^{-\frac{\mu(d+2)}{2\alpha}} \|\theta_0\|_{L^2(\mathbb{T}^2)}. \quad (5.93)$$

For short times, we use that for any equation of the form (1.1) with  $\kappa > 0$  and  $\mathbf{b}(t, \cdot)$  divergence-free and satisfying (1.3), the solution  $\theta^\kappa$  of (1.1) satisfies

$$\|\theta^\kappa\|_{C^{0,1/2}([0,\kappa^{2/\alpha}]; L^2(\mathbb{T}^d))} \leq C\kappa^{-1-\frac{d+2}{2\alpha}} \|\theta_0\|_{H^1(\mathbb{T}^d)}. \quad (5.94)$$

To get this, cover the torus  $\mathbb{T}^d$  with balls of radius  $\kappa^{1/\alpha}$  and We therefore get, using an interpolation of  $C^{0,\mu}$  between  $L^\infty$  and  $C^{0,1/2}$  and (5.71),

$$\begin{aligned} \|\theta_m - \theta_{m-1}\|_{C^{0,\mu}([0,\varepsilon_m^{2(\beta+\gamma)/(\beta-1)}]; L^2(\mathbb{T}^2))} &\leq \|\theta_m - \theta_{m-1}\|_{L^\infty([0,\varepsilon_m^{2(\beta+\gamma)/(\beta-1)}]; L^2(\mathbb{T}^2))}^{1-2\mu} \|\theta_m - \theta_{m-1}\|_{C^{0,1/2}([0,\varepsilon_m^{2(\beta+\gamma)/(\beta-1)}]; L^2(\mathbb{T}^2))}^{2\mu} \\ &\leq C\varepsilon_{m-1}^{(1-2\mu)\delta - 2q\mu(1+\frac{4}{2(\beta-1)})} \|\theta_0\|_{H^1(\mathbb{T}^2)}. \end{aligned}$$

Combining this with the above, we get

$$\begin{aligned} \|\theta_m - \theta_{m-1}\|_{C^{0,\mu}([0,1]; L^2(\mathbb{T}^2))} &\leq C \left( \varepsilon_{m-1}^{(1-\mu)\delta} (\varepsilon_m^{4(\beta+\gamma)/(\beta-1)} \wedge \varepsilon_m^{\beta+\gamma})^{-\frac{4\mu}{2\alpha}} + \varepsilon_{m-1}^{(1-2\mu)\delta - 2q\mu(1+\frac{4}{2(\beta-1)})} \right) \|\theta_0\|_{H^1(\mathbb{T}^2)}. \end{aligned}$$

For  $\mu$  small enough, depending only on  $\delta$ , we obtain (5.90).

## 5.4 The proof of Theorem 1.1

We turn to the proof of Theorem 1.1.

We first prove the theorem in the case that the initial datum  $\theta_0$  satisfies the analyticity condition (4.2). It will be convenient to define  $m_* \in \mathbb{N}$  by

$$m_* := m_{\theta_0} - 1 = \min \left\{ m \in \mathbb{N} : m \geq 2, \varepsilon_{m-1}^{1+\gamma/2} \leq R_{\theta_0} \right\} - 1. \quad (5.95)$$

We start with the observation that, due to (3.45) and the definition of  $m_*$  above, there are constants  $c_* > 0$  and  $C_* < \infty$  which depend only on  $\beta$ , such that

$$c_* R_{\theta_0}^{\frac{2q(\beta+\gamma)}{2+\gamma}} \leq c_* \varepsilon_{m_*-1}^{q(\beta+\gamma)} \leq c_* \varepsilon_{m_*}^{\beta+\gamma} \leq \kappa_{m_*} \leq C_* \varepsilon_{m_*}^{\beta+\gamma} \leq C_* R_{\theta_0}^{\frac{2(\beta+\gamma)}{2+\gamma}}. \quad (5.96)$$

Recall that  $\theta_{m_*}$  satisfies the advection-diffusion equation with diffusivity  $\kappa_{m_*}$  and vector field  $\mathbf{b}_{m_*}$ :

$$\partial_t \theta_{m_*} - \kappa_{m_*} \Delta \theta_{m_*} + \mathbf{b}_{m_*} \cdot \nabla \theta_{m_*} = 0 \quad \text{in } (0, 1) \times \mathbb{T}^2.$$

The standard  $L^2$  energy estimate and the Poincaré inequality are applicable. Since  $\langle \theta_0 \rangle = 0$ , we find that

$$\frac{d}{dt} \|\theta_{m_*}(t, \cdot)\|_{L^2(\mathbb{T}^2)}^2 = -2\kappa_{m_*} \|\nabla \theta_{m_*}(t, \cdot)\|_{L^2(\mathbb{T}^2)}^2 \leq -8\kappa_{m_*} \pi^2 \|\theta_{m_*}(t, \cdot)\|_{L^2(\mathbb{T}^2)}^2.$$

Therefore, we obtain

$$\begin{aligned} \int_0^1 \kappa_{m_*} \|\nabla \theta_{m_*}(s, \cdot)\|_{L^2(\mathbb{T}^2)}^2 ds &= \frac{1}{2} \left( \|\theta_0\|_{L^2(\mathbb{T}^2)}^2 - \|\theta_{m_*}(1, \cdot)\|_{L^2(\mathbb{T}^2)}^2 \right) \\ &\geq \|\theta_0\|_{L^2(\mathbb{T}^2)}^2 \frac{1 - e^{-8\kappa_{m_*} \pi^2}}{2} \geq 2\pi^2 \kappa_{m_*} \|\theta_0\|_{L^2(\mathbb{T}^2)}^2. \end{aligned}$$

The last inequality is valid only if  $8\kappa_{m_*} \pi^2 \leq 1$ , but this can be assumed to be valid in view of (5.95) and (5.96), by taking  $\Lambda$  larger if necessary. Therefore, we obtain

$$2\pi^2 c_* R_{\theta_0}^{\frac{2q(\beta+\gamma)}{2+\gamma}} \|\theta_0\|_{L^2(\mathbb{T}^2)}^2 \leq \kappa_{m_*} \int_0^1 \|\nabla \theta_{m_*}(s, \cdot)\|_{L^2(\mathbb{T}^2)}^2 ds. \quad (5.97)$$

By (5.72) of Proposition 5.2, for every  $m \in \{m_* + 1, \dots, M\}$ , it holds that

$$\begin{aligned} \left(1 - C\varepsilon_{m-1}^\delta\right) \kappa_{m-1} \|\nabla \theta_{m-1}\|_{L^2((0,1) \times \mathbb{T}^2)}^2 &\leq \kappa_m \|\nabla \theta_m\|_{L^2((0,1) \times \mathbb{T}^2)}^2 \\ &\leq \left(1 + C\varepsilon_{m-1}^\delta\right) \kappa_{m-1} \|\nabla \theta_{m-1}\|_{L^2((0,1) \times \mathbb{T}^2)}^2. \end{aligned} \quad (5.98)$$

We take the parameter  $\Lambda$  to be large enough that, with  $C$  as in the previous display, we have that

$$C\Lambda^{-\delta} \leq \frac{1}{100}. \quad (5.99)$$

This ensures that

$$1 - C\varepsilon_{m-1}^\delta \geq 1 - C\varepsilon_1^\delta \geq \frac{99}{100}.$$

By induction, in view of (2.9), it follows from (5.98) that

$$\min_{m \in \{m_*, \dots, M\}} \kappa_m \|\nabla \theta_m\|_{L^2((0,1) \times \mathbb{T}^2)}^2 \geq \left( \prod_{m=m_*+1}^M \left(1 - C\varepsilon_{m-1}^\delta\right) \right) \kappa_{m_*} \|\nabla \theta_{m_*}\|_{L^2((0,1) \times \mathbb{T}^2)}^2,$$

where  $M = M(\kappa)$  is an integer satisfying (3.44). Observe that, using the elementary inequality  $-2x \leq \log(1-x)$  valid for all  $x \in (0, 1/2]$ , we get

$$\log \prod_{m=m_*+1}^M \left(1 - C\varepsilon_{m-1}^\delta\right) = \sum_{m=m_*+1}^M \log \left(1 - C\varepsilon_{m-1}^\delta\right) \geq -2C \sum_{m=m_*+1}^M \varepsilon_{m-1}^\delta \geq -C\varepsilon_{m_*}^\delta.$$

Note that we used (5.99) to get the last inequality in the above display. We therefore obtain

$$\prod_{m=m_*+1}^M \left(1 - C\varepsilon_{m-1}^\delta\right) \geq \exp(-C\varepsilon_{m_*}^\delta) \geq \frac{3}{4},$$

and hence

$$\min_{m \in \{m_*, \dots, M\}} \kappa_m \|\nabla \theta_m\|_{L^2((0,1) \times \mathbb{T}^2)}^2 \geq \frac{3}{4} \kappa_{m_*} \|\nabla \theta_{m_*}\|_{L^2((0,1) \times \mathbb{T}^2)}^2.$$

By a very similar argument, using the upper bound of (5.98) rather than the lower bound, we also obtain that

$$\max_{m \in \{m_*, \dots, M\}} \kappa_m \|\nabla \theta_m\|_{L^2((0,1) \times \mathbb{T}^2)}^2 \leq \frac{4}{3} \kappa_{m_*} \|\nabla \theta_{m_*}\|_{L^2((0,1) \times \mathbb{T}^2)}^2.$$

In particular, since  $\kappa = \kappa_M$ , we have

$$\frac{3}{4} \kappa_{m_*} \|\nabla \theta_{m_*}\|_{L^2((0,1) \times \mathbb{T}^2)}^2 \leq \kappa \|\nabla \theta_M\|_{L^2((0,1) \times \mathbb{T}^2)}^2 \leq \frac{4}{3} \kappa_{m_*} \|\nabla \theta_{m_*}\|_{L^2((0,1) \times \mathbb{T}^2)}^2. \quad (5.100)$$

We can write an equation for the difference  $\theta - \theta_M$  as

$$\partial_t(\theta - \theta_M) - \kappa \Delta(\theta - \theta_M) + \mathbf{b} \cdot \nabla(\theta - \theta_M) = \nabla \cdot ((\phi - \phi_M) \nabla \theta_M) \quad \text{in } (0, 1) \times \mathbb{T}^d.$$

Recall that, by (2.39),

$$\|\phi - \phi_M\|_{L^\infty(\mathbb{R} \times \mathbb{T}^d)} \leq C\varepsilon_{M+1}^\beta = C\varepsilon_M^{\beta q}.$$

We may therefore compare  $\theta_M$  to  $\theta$  using the above displays, (3.44), (5.100) and an energy estimate:

$$\begin{aligned} & \|\theta - \theta_M\|_{L^\infty((0,1); L^2(\mathbb{T}^2))}^2 + \kappa \|\nabla \theta - \nabla \theta_M\|_{L^2((0,1) \times \mathbb{T}^2)}^2 \\ & \leq \frac{C}{\kappa} \|\phi - \phi_M\|_{L^\infty((0,1) \times \mathbb{T}^2)}^2 \|\nabla \theta_M\|_{L^2((0,1) \times \mathbb{T}^2)}^2 \\ & \leq \frac{C}{\kappa_M^2} \|\phi - \phi_M\|_{L^\infty((0,1) \times \mathbb{T}^2)}^2 \kappa_{m_*} \|\nabla \theta_{m_*}\|_{L^2((0,1) \times \mathbb{T}^2)}^2 \\ & \leq C\varepsilon_M^{-\frac{4\beta}{q+1}} \varepsilon_M^{2\beta q} \cdot \left( \kappa_{m_*} \|\nabla \theta_{m_*}\|_{L^2((0,1) \times \mathbb{T}^2)}^2 \right) \\ & = C\varepsilon_M^{2\beta(q - \frac{2}{q+1})} \left( \kappa_{m_*} \|\nabla \theta_{m_*}\|_{L^2((0,1) \times \mathbb{T}^2)}^2 \right) \\ & \leq C\Lambda^{-2M\beta(q - \frac{2}{q+1})} \left( \kappa_{m_*} \|\nabla \theta_{m_*}\|_{L^2((0,1) \times \mathbb{T}^2)}^2 \right). \end{aligned} \quad (5.101)$$

Since  $q > 1$ , the exponent of  $\Lambda$  in the last term on the right side is negative. Enlarging  $\Lambda$ , if necessary, and using that  $M \geq 1$ , we obtain

$$\|\theta - \theta_M\|_{L^\infty((0,1); L^2(\mathbb{T}^2))}^2 + \kappa \|\nabla \theta - \nabla \theta_M\|_{L^2((0,1) \times \mathbb{T}^2)}^2 \leq \frac{1}{4} \left( \kappa_{m_*} \|\nabla \theta_{m_*}\|_{L^2((0,1) \times \mathbb{T}^2)}^2 \right).$$

Combining the previous display with (5.100) and (5.97) therefore yields

$$\begin{aligned}
\kappa^{1/2} \|\nabla \theta\|_{L^2((0,1) \times \mathbb{T}^2)} &\geq \kappa^{1/2} (\|\nabla \theta_M\|_{L^2((0,1) \times \mathbb{T}^2)} - \|\nabla \theta - \nabla \theta_M\|_{L^2((0,1) \times \mathbb{T}^2)}) \\
&\geq \frac{1}{\sqrt{2}} \kappa_{m*}^{1/2} \|\nabla \theta_{m*}\|_{L^2((0,1) \times \mathbb{T}^2)} \\
&\geq \pi \sqrt{c_*} R_{\theta_0}^{\frac{q(\beta+\gamma)}{2+\gamma}} \|\theta_0\|_{L^2(\mathbb{T}^2)}. \tag{5.102}
\end{aligned}$$

This completes the proof of Theorem 1.1 in the case that the analyticity condition (4.2) is satisfied.

We now turn to the argument for general  $\theta_0 \in H^1(\mathbb{T}^2)$  of zero mean. We introduce the length scale  $L_{\theta_0}$  implicitly appearing in (1.6):

$$L_{\theta_0} := \frac{\|\theta_0\|_{L^2(\mathbb{T}^2)}}{\|\theta_0\|_{H^1(\mathbb{T}^2)}} \tag{5.103}$$

and we mollify  $\theta_0$  (viewed as a periodic function on  $\mathbb{R}^2$ ) with the standard heat kernel  $\Phi$  (with diffusion coefficient of unit size) at time  $\alpha^2 L_{\theta_0}^2$ :

$$\tilde{\theta}_0 := \theta_0 * \Phi(\alpha^2 L_{\theta_0}^2, \cdot). \tag{5.104}$$

Here  $\alpha > 0$  is a small parameter we will choose below. It is clear that  $\tilde{\theta}_0$  is a periodic and mean-zero function. It is furthermore analytic, since  $\Phi(\alpha^2 L_{\theta_0}^2, \cdot)$  is, and satisfies, for a universal constant  $C < \infty$ ,

$$\max_{|\alpha|=n} \|\partial^\alpha \tilde{\theta}_0\|_{L^2(\mathbb{T}^2)} \leq \|\theta_0\|_{L^2(\mathbb{T}^2)} n! \left( \frac{C}{\alpha L_{\theta_0}} \right)^n, \quad \forall n \in \mathbb{N}.$$

That is, the analyticity condition (4.2) is valid for  $R_{\tilde{\theta}_0} = c\alpha L_{\theta_0}$ . If we let  $\tilde{\theta}$  be the solution of (1.1) with  $\tilde{\theta}_0$  in place of  $\theta_0$ , then  $\theta - \tilde{\theta}$  is also a solution of the same equation with initial data  $\theta_0 - \tilde{\theta}_0$ , and therefore the incompressibility of  $\mathbf{b}$  implies as in (1.5) that

$$\|(\theta - \tilde{\theta})(t, \cdot)\|_{L^2(\mathbb{T}^2)}^2 + 2\kappa \|\nabla(\theta - \tilde{\theta})\|_{L^2((0,t) \times \mathbb{T}^2)}^2 = \|\theta_0 - \tilde{\theta}_0\|_{L^2(\mathbb{T}^2)}^2, \quad \forall t > 0. \tag{5.105}$$

In view of (5.102) (applied to  $\tilde{\theta}$  replacing  $\theta$  and  $\tilde{\theta}_0$  replacing  $\theta_0$ ) and the triangle inequality, it suffices to show that we can choose  $\alpha$  in such a way that

$$\|\theta_0 - \tilde{\theta}_0\|_{L^2(\mathbb{T}^2)} \leq \frac{1}{2} \pi \sqrt{c_*} R_{\tilde{\theta}_0}^{\frac{q(\beta+\gamma)}{2+\gamma}} \|\theta_0\|_{L^2(\mathbb{T}^2)}. \tag{5.106}$$

We next compute

$$\|\tilde{\theta}_0\|_{L^2(\mathbb{T}^2)}^2 = \|\theta_0\|_{L^2(\mathbb{T}^2)}^2 - 2 \int_0^{\alpha^2 L_{\theta_0}^2} \int_{\mathbb{T}^2} |(\theta_0 * \nabla \Phi(s, \cdot))(x)|^2 dx dt.$$

Using that

$$\int_0^{\alpha^2 L_{\theta_0}^2} \int_{\mathbb{T}^2} |(\theta_0 * \nabla \Phi(s, \cdot))(x)|^2 dx dt \leq \alpha^2 L_{\theta_0}^2 \|\nabla \theta_0\|_{L^2(\mathbb{T}^2)}^2 \leq C \alpha^2 \|\theta_0\|_{L^2(\mathbb{T}^2)}^2,$$

we therefore obtain

$$(1 - C\alpha) \|\theta_0\|_{L^2(\mathbb{T}^2)} \leq \|\tilde{\theta}_0\|_{L^2(\mathbb{T}^2)} \leq \|\theta_0\|_{L^2(\mathbb{T}^2)}. \tag{5.107}$$

Since  $R_{\tilde{\theta}_0} = c\alpha L_{\theta_0}$ , the inequality (5.106) will be valid provided that  $\alpha$  satisfies

$$\alpha \leq cL_{\theta_0}^{\frac{q(\beta+\gamma)}{2+\gamma-q(\beta+\gamma)}}.$$

With this choice of  $\alpha$ , we observe that (5.102), (5.105), and (5.106) imply

$$\kappa^{1/2} \|\nabla \theta\|_{L^2((0,1) \times \mathbb{T}^2)} \geq \frac{1}{2} \pi \sqrt{c_*} R_{\theta_0}^{\frac{q(\beta+\gamma)}{2+\gamma}} \|\theta_0\|_{L^2(\mathbb{T}^2)} = cL_{\theta_0}^{\frac{q(\beta+\gamma)}{2+\gamma-q(\beta+\gamma)}} \|\theta_0\|_{L^2(\mathbb{T}^2)}.$$

This completes the proof of the theorem.

As a final remark, we note that the exponent of  $L_{\theta_0}$ , namely  $\frac{q(\beta+\gamma)}{2+\gamma-q(\beta+\gamma)}$ , can be taken arbitrarily close to  $\frac{\beta}{2-\beta}$ , by taking  $q$  closer to 1 than in (2.2), at the cost of all constants depending additionally on  $q$ . With  $\alpha = \beta - 1$ , this matches what was promised in (1.6).  $\square$

**Remark 5.4** (Uniform continuity in time of the scalar variance). By using (5.90) and repeating the argument of Remark 5.3 to get a similar bound for the difference  $\theta - \theta_M$ , using in particular (5.92), (5.94) and (5.101), we obtain

$$\begin{aligned} \|\theta\|_{C^{0,\mu}([0,1];L^2(\mathbb{T}^2))} &\leq \|\theta_{m_*}\|_{C^{0,\mu}([0,1];L^2(\mathbb{T}^2))} + \|\theta - \theta_M\|_{C^{0,\mu}([0,1];L^2(\mathbb{T}^2))} + \sum_{m=m_*}^M \|\theta_m - \theta_{m-1}\|_{C^{0,\mu}([0,1];L^2(\mathbb{T}^2))} \\ &\leq C_{\theta_0}, \end{aligned} \quad (5.108)$$

where  $C_{\theta_0}$  depends only on  $\|\theta_0\|_{H^1(\mathbb{T}^2)}$  and  $\beta$ .

## 5.5 Lack of selection in the vanishing diffusivity limit

The goal of this section is to prove the following statement concerning the lack of selection principle for vanishing diffusivity limits of the advection-diffusion equation to the transport equation. Our result applies to a large class of  $\dot{H}^2(\mathbb{T}^2)$  initial conditions, but not to all of them.

**Proposition 5.5.** *Fix  $\beta \in [68/67, 4/3]$ . There exists a constant  $C_* = C_*(\beta) \geq 1$  such that the following holds. For every parameters  $A \in (0, 1]$  and  $B > 1$ , assume that  $\Lambda$  is taken sufficiently large with respect to  $\beta, A, B$  to ensure that (see condition (5.138) below):*

$$C_* \Lambda^{-\delta} \leq \min \left\{ A, B^{-\frac{2}{2+\gamma}(2-\beta+\frac{2\beta}{q+1})} \right\}.$$

Fix the sequence  $\{\varepsilon_m\}_{m \geq 0}$  according to (2.8). Choose an initial datum  $\theta_0 \in \dot{H}^2(\mathbb{T}^2)$ . Define the length scale  $L_{\theta_0} := \frac{\|\theta_0\|_{L^2(\mathbb{T}^2)}}{\|\nabla \theta_0\|_{L^2(\mathbb{T}^2)}}$  and let  $m_* \geq 1$  be the unique integer such that  $\varepsilon_{m_*} \leq C_* L_{\theta_0}^{\frac{2}{2+\gamma-2q\delta}} < \varepsilon_{m_*-1}$ . Assuming that  $\theta_0$  satisfies (see conditions (5.139) and (5.140) below):

$$\frac{\|\nabla \theta_0\|_{L^2(\mathbb{T}^2)}^4}{\|\theta_0\|_{L^2(\mathbb{T}^2)}^2 \|\Delta \theta_0\|_{L^2(\mathbb{T}^2)}^2} \geq A, \quad \text{and} \quad C_* L_{\theta_0}^{\frac{2}{2+\gamma-2q\delta}} \leq B \varepsilon_{m_*},$$

there exist two sequences of diffusivities  $\{\kappa_m^{(1)}\}_{m \in \mathbb{N}}$  and  $\{\kappa_m^{(2)}\}_{m \in \mathbb{N}}$ , both converging to 0 as  $m \rightarrow \infty$ , such that the corresponding solutions  $\{\theta^{\kappa_m^{(1)}}\}_m$  and  $\{\theta^{\kappa_m^{(2)}}\}_m$  of the advection-diffusion equation with initial data  $\theta_0$  and drift  $\mathbf{b}$ , converge as  $m \rightarrow \infty$  in  $C^{0,\mu}((0, 1); L^2(\mathbb{T}^2))$  (for some  $\mu > 0$ ) to two distinct weak solutions of the associated transport equation.

In order to prove Proposition 5.5, we are going to find two sequences  $\{\kappa_m^{(1)}\}_{m \in \mathbb{N}}$  and  $\{\kappa_m^{(2)}\}_{m \in \mathbb{N}}$  such that:

- $\kappa_m^{(1)}$  and  $\kappa_m^{(2)}$  belong to the  $m$ th interval in (3.43), namely  $[\frac{1}{2}\varepsilon_m^{\frac{2\beta}{q+1}}, 2\varepsilon_m^{\frac{2\beta}{q+1}}]$ ;
- The choice is made such that the ratio of  $\kappa_m^{(1)}$  to  $\kappa_m^{(2)}$  is not close to one in the sense that it lies in  $[\frac{1}{2}, 2] \setminus [\frac{3}{4}, \frac{4}{3}]$ .
- Each of these infinite sequences satisfies the required recurrence formula for the effective diffusivities, namely (5.109) below.

The existence of this pair of sequences is a consequence of Lemma 5.6, below. We then turn our attention to the two corresponding solutions of the advection-diffusion equation (4.3) with  $m = m_*$ , with  $m_*$  defined as above in (5.95), which we denote by  $\theta_{m_*}^{(1)}$  and  $\theta_{m_*}^{(2)}$ . We show in Lemma 5.7 that these two solutions have to be different, at least for short times  $t$ , as they experience a different amount of diffusion (since  $\kappa_{m_*}^{(1)} \neq \kappa_{m_*}^{(2)}$ ). Finally, we appeal to the bound (5.71) proved above to obtain a positive lower bound on the size of  $\|\theta_m^{(1)} - \theta_m^{(2)}\|_{L^2}$  for all  $m \geq m_*$ . The estimate (5.71) also gives us the convergence in  $L_t^\infty L_x^2$  of  $\theta_m^{(i)}$  to a solution of the transport equation, which is evidently different for  $i = 1$  and  $i = 2$ .

We begin with the following lemma, which revisits the analysis of the recurrence (3.42) given in Lemma 3.4 and extracts some extra information.

**Lemma 5.6.** *There exists  $C(\beta) < \infty$  and  $\rho(\beta) > 0$  such that, if  $\Lambda \geq C$ , then, for each  $\tau \in [\frac{1}{2}, 2]$ , there exists an infinite sequence  $\{\kappa_m\}_{m \in \mathbb{N}}$  satisfying*

$$\kappa_{m-1} = \bar{\mathbf{K}}_m^{\kappa_m}, \quad \forall m \geq 1, \quad (5.109)$$

such that

$$\left| \frac{\kappa_m \cdot \frac{4}{3}\sqrt{5}}{\varepsilon_m^{\frac{2\beta}{q+1}}} - \tau^{(-1)^m} \right| \leq C\varepsilon_m^\rho \leq \frac{1}{100}, \quad \forall m \in \mathbb{N}. \quad (5.110)$$

*Proof.* We first suppose that  $M \in \mathbb{N}$  and  $\{\kappa_0, \kappa_1, \dots, \kappa_M\} \subseteq (0, \infty)$  is a sequence of positive numbers satisfying, for every  $m \in \{1, \dots, M\}$  the recurrence relation (5.109). We also assume that

$$\tau := \frac{\kappa_M \cdot \frac{4}{3}\sqrt{5}}{\varepsilon_M^{\frac{2\beta}{q+1}}} \in \left[ \frac{1}{2}, 2 \right]. \quad (5.111)$$

Using the recurrence relation (3.53) for the effective diffusivities, expressed in terms of the quantity  $s_m$  defined in the proof of Lemma 3.4 by

$$s_m := \frac{\kappa'_m \cdot \frac{4}{3}\sqrt{5}}{a_m \varepsilon_m^{2+\gamma}},$$

and  $\kappa'_m$  is defined in (3.47). Recall that  $a_m$  and the exponents  $\beta$ ,  $\gamma$  and  $q$  are defined in such a way that  $a_m \varepsilon_m^{2+\gamma} = \varepsilon_m^{\beta+\gamma} = \varepsilon_m^{\frac{2\beta}{q+1}}$ . The recurrence relation for  $s_m$  appeared above in (3.51), and in view of the identity in (3.52) it can be written as

$$s_{m-1} = \frac{1}{s_m} + \underbrace{\varepsilon_m^{2\gamma} s_m}_{\leq C\varepsilon_m^{2\gamma}} + \underbrace{\left( \left( \frac{\varepsilon_m}{\varepsilon_{m-1}} \right)^\beta - 1 \right) s_m}_{\leq C\varepsilon_m^{1/q}} \underbrace{\left( \varepsilon_m^{2\gamma} + \frac{1}{s_m^2} \right)}_{\leq C}, \quad (5.112)$$

where the bound for the three terms on the right side come from (3.54) and (2.10). We therefore obtain, for every  $m \in \{1, \dots, M\}$ ,

$$\left| s_{m-1} - \frac{1}{s_m} \right| \leq C \varepsilon_m^{2\gamma \wedge \frac{1}{q}}.$$

Iterating this inequality yields, for every  $n \in \{0, \dots, M\}$ ,

$$\left| s_{M-n}^{(-1)^n} - s_M \right| \leq C \varepsilon_{M-n+1}^{2\gamma \wedge \frac{1}{q}}. \quad (5.113)$$

If we enlarge  $\Lambda$  so that  $C_{(5.113)} \varepsilon_1^{2\gamma \wedge \frac{1}{q}} \leq \frac{1}{500}$ , then we obtain,

$$\max_{n \in \{0, \dots, M\}} \left| s_{M-n}^{(-1)^n} - s_M \right| \leq C \varepsilon_{M-n+1}^{2\gamma \wedge \frac{1}{q}} \leq \frac{1}{500}.$$

By enlarging  $\Lambda$ , if necessary, we may use (3.58) to obtain, for every  $n \in \{0, \dots, M-1\}$ ,

$$\max_{m \in \{n, \dots, M-1\}} \max \left\{ \frac{\kappa_m}{\kappa'_m} - 1, \frac{\kappa'_m}{\kappa_m} - 1 \right\} \leq C \varepsilon_{n+1}^{\delta \wedge \gamma} \leq \frac{1}{500}.$$

Combining these yields, some positive exponent  $\rho(\beta) > 0$ ,

$$\left| \frac{\kappa_n \cdot \frac{4}{3}\sqrt{5}}{\varepsilon_n^{\frac{2\beta}{q+1}}} - \tau^{(-1)^{M-n}} \right| \leq C \varepsilon_n^\rho \leq \frac{1}{100}. \quad (5.114)$$

Finally, the existence of an infinite sequence satisfying (5.110) is obtained by recursively defining, for each  $M \in \mathbb{N}$ , a finite sequence  $\{\kappa_m^{(M)}\}_{0 \leq m \leq M}$  by

$$\begin{cases} \kappa_{m-1}^{(M)} = \bar{\mathbf{K}}_m^{\kappa_m^{(M)}}, & m \in \{1, \dots, M\}, \\ \kappa_M^{(M)} := \tau^{(-1)^M} (\frac{4}{3}\sqrt{5})^{-1} \varepsilon_M^{\frac{2\beta}{q+1}}. \end{cases} \quad (5.115)$$

Sending  $M \rightarrow \infty$ , we find that  $\kappa_m^{(M)}$  converges, for each fixed  $m$ , yielding the desired infinite sequence.  $\square$

We next argue that the solutions of two advection-diffusion equations with the same drift but different diffusivities must be different, at least for short times. In the following lemma,  $\mathbf{u}$  plays the role of  $\mathbf{b}_{m*}$  and  $\kappa^{(i)}$  plays the role of  $\kappa_{m*}^{(i)}$ .

**Lemma 5.7.** *Suppose that  $0 < \kappa^{(1)} \leq \frac{4}{5}\kappa^{(2)}$  and  $\mathbf{u} : \mathbb{T}^2 \rightarrow \mathbb{R}^2$  is a smooth, incompressible vector field. Let  $\theta_0 \in H^2(\mathbb{T}^2)$  and, for each  $i \in \{1, 2\}$ , let  $\theta^{(i)}$  be the solution of*

$$\begin{cases} \partial_t \theta^{(i)} - \kappa^{(i)} \Delta \theta^{(i)} + \mathbf{u} \cdot \nabla \theta^{(i)} = 0 & \text{in } (0, \infty) \times \mathbb{R}^2, \\ \theta^{(i)} = \theta_0 & \text{on } \{0\} \times \mathbb{R}^2. \end{cases} \quad (5.116)$$

Then, for every time  $t$  satisfying

$$0 \leq t \leq \frac{1}{50} \min \left\{ \frac{1}{\|\nabla \mathbf{u}\|_{L^\infty([0,1] \times \mathbb{T}^2)}}, \frac{\|\nabla \theta_0\|_{L^2(\mathbb{T}^2)}^2}{\kappa^{(1)} \|\Delta \theta_0\|_{L^2(\mathbb{T}^2)}^2} \right\},$$

we have the estimate

$$\|\theta^{(1)}(t, \cdot)\|_{L^2(\mathbb{T}^2)}^2 - \|\theta^{(2)}(t, \cdot)\|_{L^2(\mathbb{T}^2)}^2 \geq \frac{11}{40} \kappa^{(1)} t \|\nabla \theta_0\|_{L^2(\mathbb{T}^2)}^2. \quad (5.117)$$

*Proof.* With  $\mathbf{u}$  smooth and incompressible, consider, for  $\kappa > 0$ , the drift-diffusion equation

$$\partial_t \theta - \kappa \Delta \theta + \mathbf{u} \cdot \nabla \theta = 0, \quad \theta|_{t=0} = \theta_0, \quad (5.118)$$

where  $\theta_0 \in H^2(\mathbb{T}^2)$  has zero mean. Denote  $L := \|\nabla \mathbf{u}\|_{L^\infty([0,1] \times \mathbb{T}^2)}$ . Standard energy estimates (in each of the space  $L^2$ ,  $H^1$ , and  $H^2$ ) yields, for  $t \geq 0$ ,

$$\begin{aligned} & \left| \|\theta_0\|_{L^2(\mathbb{T}^2)}^2 - \|\theta(t, \cdot)\|_{L^2(\mathbb{T}^2)}^2 - 2\kappa t \|\nabla \theta_0\|_{L^2(\mathbb{T}^2)}^2 \right| \\ & \leq 2\kappa t \|\nabla \theta_0\|_{L^2(\mathbb{T}^2)}^2 (tL)^2 \exp(2tL) + 2\kappa t \|\nabla \theta_0\|_{L^2(\mathbb{T}^2)}^2 (tL) \exp(2tL) + 2\kappa^2 t^2 \|\Delta \theta_0\|_{L^2(\mathbb{T}^2)}^2 \exp(2tL). \end{aligned}$$

As a consequence, we obtain

$$\left| \frac{\|\theta_0\|_{L^2(\mathbb{T}^2)}^2 - \|\theta(t, \cdot)\|_{L^2(\mathbb{T}^2)}^2}{2\kappa t \|\nabla \theta_0\|_{L^2}^2} - 1 \right| \leq \left( tL + (tL)^2 + \kappa t \frac{\|\Delta \theta_0\|_{L^2(\mathbb{T}^2)}^2}{\|\nabla \theta_0\|_{L^2(\mathbb{T}^2)}^2} \right) \exp(2tL). \quad (5.119)$$

Therefore, for  $t$  satisfying

$$t \leq \frac{1}{50} \min \left\{ \frac{1}{L}, \frac{\|\nabla \theta_0\|_{L^2(\mathbb{T}^2)}^2}{\kappa \|\Delta \theta_0\|_{L^2(\mathbb{T}^2)}^2} \right\}, \quad (5.120)$$

we obtain

$$\frac{19}{10} \kappa t \|\nabla \theta_0\|_{L^2(\mathbb{T}^2)}^2 \leq \|\theta_0\|_{L^2(\mathbb{T}^2)}^2 - \|\theta(t, \cdot)\|_{L^2(\mathbb{T}^2)}^2 \leq \frac{21}{10} \kappa t \|\nabla \theta_0\|_{L^2(\mathbb{T}^2)}^2. \quad (5.121)$$

Now consider the solutions  $\theta^{(1)}$  and  $\theta^{(2)}$  of (5.116) with diffusivity parameters  $\kappa^{(1)}, \kappa^{(2)} > 0$  satisfying  $\kappa^{(2)} \geq \frac{5}{4} \kappa^{(1)}$ . By (5.119) and the triangle inequality, for all  $t \geq 0$  satisfying (5.120) with  $\kappa = \kappa^{(1)}$ , we obtain

$$\|\theta^{(1)}(t, \cdot)\|_{L^2(\mathbb{T}^2)}^2 - \|\theta^{(2)}(t, \cdot)\|_{L^2(\mathbb{T}^2)}^2 \geq \left( \frac{19}{10} \kappa^{(2)} - \frac{21}{10} \kappa^{(1)} \right) t \|\nabla \theta_0\|_{L^2}^2.$$

Since  $\kappa^{(2)} \geq \frac{5}{4} \kappa^{(1)}$ , we obtain (5.117). The proof is complete.  $\square$

We next argue that the previous two lemmas imply Proposition 5.5.

*Proof of Proposition 5.5.* We apply Lemma 5.6 twice, with  $\tau = \tau_1 := 1/2$  and  $\tau = \tau_2 := 3/4$ , to find two sequences  $\{\kappa_m^{(1)}\}_{m \in \mathbb{N}}$  and  $\{\kappa_m^{(2)}\}_{m \in \mathbb{N}}$  satisfying the recurrence (5.109) and, for each  $i \in \{1, 2\}$ ,

$$\left| \frac{\kappa_n^{(i)} \cdot \frac{4}{3} \sqrt{5}}{\varepsilon_n^{\frac{2\beta}{q+1}}} - \tau_i^{(-1)^n} \right| \leq C \varepsilon_n^\rho \leq \frac{1}{100}, \quad \forall n \in \mathbb{N}. \quad (5.122)$$

This implies in particular that, for every  $n \in \mathbb{N}$ ,

$$\frac{5}{4} \leq \max \left\{ \frac{\kappa_n^{(1)}}{\kappa_n^{(2)}}, \frac{\kappa_n^{(2)}}{\kappa_n^{(1)}} \right\} \leq 2. \quad (5.123)$$

According to the two sequences  $\{\kappa_m^{(1)}\}_{m \in \mathbb{N}}$  and  $\{\kappa_m^{(2)}\}_{m \in \mathbb{N}}$  constructed above, we denote by  $\theta_m^{(1)}$  and  $\theta_m^{(2)}$  the solutions of the initial-value problems

$$\begin{cases} \partial_t \theta_m^{(i)} - \kappa_m^{(i)} \Delta \theta_m^{(i)} + \mathbf{b}_m \cdot \nabla \theta_m^{(i)} = 0 & \text{in } (0, \infty) \times \mathbb{R}^2, \\ \theta_m^{(i)} = \theta_0 & \text{on } \{0\} \times \mathbb{R}^2, \end{cases} \quad (5.124)$$

where  $\mathbf{b}_m$  is as in the rest of the paper (see (2.36)).

Our next goal is to show that if  $\theta_0 \in H^2(\mathbb{T}^2)$  has zero mean and satisfies certain conditions (see (5.139) and (5.140) below), then there exists a time  $t \in (0, 1]$  such that  $|\|\theta_m^{(1)}(t, \cdot)\|_{L^2(\mathbb{T}^2)}^2 - \|\theta_m^{(2)}(t, \cdot)\|_{L^2(\mathbb{T}^2)}^2|$  is bounded from below by a positive quantity, uniformly in  $m$  (for  $m$  sufficiently large). In order to achieve this uniform in  $m$  lower bound, as in the proof of Theorem 1.1 let us temporarily assume that  $\theta_0$  is an analytic function, satisfying the quantitative estimate (4.2) for some  $R_{\theta_0} > 0$ . (At the end of the argument, we approximate the given  $H^2$  data with an analytic one, akin to (5.104).) For  $\theta_0$  satisfying (4.2), let  $m_* = m_*(R_{\theta_0}, \Lambda, \beta) \in \mathbb{N}$  be defined according to (5.95); we also record the bound (5.96) which relates  $\kappa_{m_*}$  and  $\varepsilon_{m_*}$  to  $R_{\theta_0}$ .

Applying (5.117) to the problem (5.124) with  $m = m_* \in \mathbb{N}$  (as defined in (5.95)), and appealing to the  $\mathbf{b}_{m_*}$  estimate in (2.72), to the bound for  $\kappa_{m_*}^{(1)}$  implied by (5.122) and (5.123), we obtain for every  $t \in (0, 1]$  satisfying

$$0 < t \leq \frac{1}{300} \min \left\{ 2^{-22} \varepsilon_{m_*}^{2-\beta}, \frac{\|\nabla \theta_0\|_{L^2(\mathbb{T}^2)}^2}{\|\Delta \theta_0\|_{L^2(\mathbb{T}^2)}^2} \varepsilon_{m_*}^{-\frac{2\beta}{q+1}} \right\}, \quad (5.125)$$

the estimate

$$\left| \|\theta_{m_*}^{(1)}(t, \cdot)\|_{L^2(\mathbb{T}^2)}^2 - \|\theta_{m_*}^{(2)}(t, \cdot)\|_{L^2(\mathbb{T}^2)}^2 \right| \geq \frac{11}{80} \kappa_{m_*}^{(1)} t \|\nabla \theta_0\|_{L^2(\mathbb{T}^2)}^2.$$

In light of the energy identity for (5.124), the above estimate yields a bound without squares

$$\left| \|\theta_{m_*}^{(1)}(t, \cdot)\|_{L^2(\mathbb{T}^2)} - \|\theta_{m_*}^{(2)}(t, \cdot)\|_{L^2(\mathbb{T}^2)} \right| \geq \frac{11}{160} \frac{\kappa_{m_*}^{(1)} t \|\nabla \theta_0\|_{L^2(\mathbb{T}^2)}^2}{\|\theta_0\|_{L^2(\mathbb{T}^2)}}. \quad (5.126)$$

In order to obtain an estimate which holds for all  $m \geq m_*$ , we appeal to (5.71) (applicable since (5.95) was designed to match (5.70)) and the triangle inequality, to arrive at:

$$\begin{aligned} & \left| \|\theta_m^{(1)}(t, \cdot)\|_{L^2(\mathbb{T}^2)} - \|\theta_m^{(2)}(t, \cdot)\|_{L^2(\mathbb{T}^2)} \right| \\ & \geq \left| \|\theta_{m_*}^{(1)}(t, \cdot)\|_{L^2(\mathbb{T}^2)} - \|\theta_{m_*}^{(2)}(t, \cdot)\|_{L^2(\mathbb{T}^2)} \right| - 2 \sup_{i \in \{1, 2\}} \sup_{m \geq m_*} \|\theta_m^{(i)} - \theta_{m_*}^{(i)}\|_{L^\infty((0, 1); L^2(\mathbb{T}^2))} \\ & \geq \frac{11}{160} \frac{\kappa_{m_*}^{(1)} t \|\nabla \theta_0\|_{L^2(\mathbb{T}^2)}^2}{\|\theta_0\|_{L^2(\mathbb{T}^2)}} - C \varepsilon_{m_*}^\delta \|\theta_0\|_{L^2(\mathbb{T}^2)}. \end{aligned} \quad (5.127)$$

The constant  $C_{(5.127)}$  may be computed from  $C_{(5.71)}$ , and the  $C = C(\beta)$  which appears in the bound  $\sum_{m \geq m_*} \varepsilon_m^\delta \leq C \varepsilon_{m_*}^\delta$ . We emphasize that (5.127) is valid for every  $t$  satisfying (5.125). We need the right side of (5.127) to be positive, which together with the bound for  $\kappa_{m_*}^{(1)}$  implied by (5.122) and (5.123), translates into the requirement that  $t$  satisfies

$$t \geq 10^3 C_{(5.127)} \varepsilon_{m_*}^\delta \varepsilon_{m_*}^{-\frac{2\beta}{q+1}} \frac{\|\theta_0\|_{L^2(\mathbb{T}^2)}^2}{\|\nabla \theta_0\|_{L^2(\mathbb{T}^2)}^2}. \quad (5.128)$$

Combining the above estimates, we arrive at the conclusion that if the initial datum  $\theta_0$  and the parameter  $\Lambda$  are chosen such that the inequality

$$10^3 C_{(5.127)} \varepsilon_{m_*}^\delta \varepsilon_{m_*}^{-\frac{2\beta}{q+1}} \frac{\|\theta_0\|_{L^2(\mathbb{T}^2)}^2}{\|\nabla \theta_0\|_{L^2(\mathbb{T}^2)}^2} \leq \frac{1}{400} \min \left\{ 2^{-22} \varepsilon_{m_*}^{2-\beta}, \frac{\|\nabla \theta_0\|_{L^2(\mathbb{T}^2)}^2}{\|\Delta \theta_0\|_{L^2(\mathbb{T}^2)}^2} \varepsilon_{m_*}^{-\frac{2\beta}{q+1}} \right\} \quad (5.129)$$

holds, then there exists an interval of times  $t \in [0, 1]$  which satisfies both (5.125) and (5.128), and moreover, for all such times  $t$  by (5.127) we have the bound

$$\begin{aligned} |\|\theta_m^{(1)}(t, \cdot)\|_{L^2(\mathbb{T}^2)} - \|\theta_m^{(2)}(t, \cdot)\|_{L^2(\mathbb{T}^2)}| &\geq \frac{1}{16} \frac{\kappa_{m_*}^{(1)} t \|\nabla \theta_0\|_{L^2(\mathbb{T}^2)}^2}{\|\theta_0\|_{L^2(\mathbb{T}^2)}} \\ &\geq C_{(5.127)} \varepsilon_{m_*}^\delta \|\theta_0\|_{L^2(\mathbb{T}^2)} \\ &\geq C_{(5.127)} R_{\theta_0}^{\frac{2q\delta}{2+\gamma}} \|\theta_0\|_{L^2(\mathbb{T}^2)}, \quad \forall m \geq m_*. \end{aligned} \quad (5.130)$$

In the second inequality we have appealed to (5.128), while in the third inequality we have used (5.96), with  $R_{\theta_0}$  being the analyticity radius of the initial data  $\theta_0$ .

It thus remains to augment the condition (5.129) and the inequality (5.130) with an approximation argument, which replaces an  $H^2(\mathbb{T}^2)$  function of zero mean with an analytic one. For a given  $\theta_0 \in H^2(\mathbb{T}^2)$ , we proceed as in the proof of Theorem 1.1. We define the length scale  $L_{\theta_0}$  according to (5.103), and we mollify  $\theta_0$  with the standard heat kernel (with diffusion coefficient of unit size) at time  $\alpha^2 L_{\theta_0}^2$ , with  $\alpha > 0$  to be determined later in the proof; the resulting analytic function  $\tilde{\theta}_0$  defined in (5.104) satisfies the analyticity condition (4.2) for  $R_{\tilde{\theta}_0} = c\alpha L_{\theta_0}$ , for a universal constant  $c \in (0, 1]$ . For  $m \in \mathbb{N}$  and  $i \in \{1, 2\}$ , denote by  $\theta_m^{(i)}$  the solution of (5.124) with initial  $\theta_0$ , and by  $\tilde{\theta}_m^{(i)}$  the solution of (5.124) with initial  $\tilde{\theta}_0$ . Using linearity, the energy balance (5.105), the triangle inequality, and (5.130) we thus deduce that there exists  $t \in [0, 1]$  such that

$$\begin{aligned} |\|\theta_m^{(1)}(t, \cdot)\|_{L^2(\mathbb{T}^2)} - \|\theta_m^{(2)}(t, \cdot)\|_{L^2(\mathbb{T}^2)}| &\geq |\|\tilde{\theta}_m^{(1)}(t, \cdot)\|_{L^2(\mathbb{T}^2)} - \|\tilde{\theta}_m^{(2)}(t, \cdot)\|_{L^2(\mathbb{T}^2)}| - 2\|\theta_0 - \tilde{\theta}_0\|_{L^2(\mathbb{T}^2)} \\ &\geq C_{(5.127)} R_{\tilde{\theta}_0}^{\frac{2q\delta}{2+\gamma}} \|\tilde{\theta}_0\|_{L^2(\mathbb{T}^2)} - 2\|\theta_0 - \tilde{\theta}_0\|_{L^2(\mathbb{T}^2)}, \end{aligned} \quad (5.131)$$

for all  $m \geq m_*$ , where  $m_*$  is defined according to (5.95), with  $R_{\theta_0}$  replaced by  $R_{\tilde{\theta}_0}$ . Naturally, the existence of the time  $t$  in (5.131) requires that (5.129) holds, with  $\theta_0$  replaced by  $\tilde{\theta}_0$ . Since  $\theta_0$  and its mollified version  $\tilde{\theta}_0$  are close in  $L^2(\mathbb{T}^2)$ , see estimate (5.107), and recalling that  $R_{\tilde{\theta}_0} = \alpha L_{\theta_0}/C$ , we deduce from (5.131) that for some universal constant  $C \geq 1$ , we have

$$|\|\theta_m^{(1)}(t, \cdot)\|_{L^2(\mathbb{T}^2)} - \|\theta_m^{(2)}(t, \cdot)\|_{L^2(\mathbb{T}^2)}| \geq \frac{1}{2} C_{(5.127)} (C^{-1} \alpha L_{\theta_0})^{\frac{2q\delta}{2+\gamma}} \|\theta_0\|_{L^2(\mathbb{T}^2)} - C\alpha \|\theta_0\|_{L^2(\mathbb{T}^2)}, \quad (5.132)$$

for all  $m \geq m_*$ . Optimizing the above estimate with respect to  $\alpha$ , we arrive at

$$\alpha = C^{-1} L_{\theta_0}^{\frac{2q\delta}{2+\gamma-2q\delta}} \quad (5.133)$$

where  $C = C(\beta) \geq 1$  is a computable constant. Here we used that the parameters  $q, \gamma, \delta$  are merely functions of  $\beta$  (see (2.2)–(2.7)), and they satisfy  $2 + \gamma - 2q\delta > 0$  for all  $\beta \in (1, 4/3)$ . Having chosen the parameter  $\alpha$ , we finally deduce from (5.132) that for some  $C = C(\beta) > 0$ , there exists  $t \in [0, 1]$  such that

$$|\|\theta_m^{(1)}(t, \cdot)\|_{L^2(\mathbb{T}^2)} - \|\theta_m^{(2)}(t, \cdot)\|_{L^2(\mathbb{T}^2)}| \geq C^{-1} L_{\theta_0}^{\frac{2q\delta}{2+\gamma-2q\delta}} \|\theta_0\|_{L^2(\mathbb{T}^2)}, \quad (5.134)$$

for all  $m \geq m_*$ . Importantly, the above estimate is valid assuming that (5.129) holds with  $\theta_0$  replaced by  $\tilde{\theta}_0$ .

In order to conclude, we identify for which  $\theta_0 \in H^2(\mathbb{T}^2)$  the condition (5.129) — with  $\theta_0$  replaced by  $\tilde{\theta}_0$  — is non-vacuous. Recall, by definition, that  $\tilde{\theta}_0$  is obtained from  $\theta_0$  by running the heat equation (at unit diffusivity) forward in time, up to the small time  $\alpha^2 L_{\theta_0}^2$ . Since  $L_0 \leq (2\pi)^{-1}$  (recall (5.103)) and since the constant  $C$  appearing in (5.133) may be increased by a factor of 10 if needed,

we thus have that  $2^{-1}\|\theta_0\|_{\dot{H}^k(\mathbb{T}^2)} \leq \|\tilde{\theta}_0\|_{\dot{H}^k(\mathbb{T}^2)} \leq \|\theta_0\|_{\dot{H}^k(\mathbb{T}^2)}$ , for  $k \in \{0, 1, 2\}$ . Therefore, if we are willing to give up a factor of 16, the norms  $\|\tilde{\theta}_0\|_{\dot{H}^k(\mathbb{T}^2)}$  appearing in (5.129) for  $k \in \{0, 1, 2\}$  may be replaced by the corresponding  $\|\theta_0\|_{\dot{H}^k(\mathbb{T}^2)}$  norms. With this in mind, we now summarize condition (5.129) as requiring two bounds:

$$(C'\varepsilon_{m_*}^\delta)^{\beta-2-\frac{2\beta}{q+1}} \leq \frac{\|\nabla\theta_0\|_{L^2(\mathbb{T}^2)}^2}{\|\theta_0\|_{L^2(\mathbb{T}^2)}^2}, \quad (5.135)$$

$$C'\varepsilon_{m_*}^\delta \leq \frac{\|\nabla\theta_0\|_{L^2(\mathbb{T}^2)}^4}{\|\theta_0\|_{L^2(\mathbb{T}^2)}^2 \|\Delta\theta_0\|_{L^2(\mathbb{T}^2)}^2}, \quad (5.136)$$

for some sufficiently large  $C' = C'(\beta) \geq 1$ . In the above two inequalities, we recall from (5.95), from the definition  $R_{\tilde{\theta}_0} = \alpha L_{\theta_0}/C$ , and from the choice of  $\alpha$  in (5.133), that  $m_* \geq 1$  is defined as the integer which satisfies

$$\varepsilon_{m_*} \leq R_{\tilde{\theta}_0}^{\frac{2}{2+\gamma}} < \varepsilon_{m_*-1}, \quad \text{where} \quad R_{\tilde{\theta}_0} = (C'')^{-1} L_{\theta_0}^{\frac{2+\gamma}{2+\gamma-2q\delta}} \quad (5.137)$$

and  $C'' = C''(\beta) \geq 1$  is sufficiently large.

At this stage, as in the statement of the Proposition, fix two parameters  $A \in (0, 1]$  and  $B > 1$ . Then, choose  $\Lambda$  large enough, solely in terms of  $A$ ,  $B$ , and  $\beta$ , to ensure that

$$C'\Lambda^{-\delta} \leq A, \quad (C'\Lambda^{-\delta})(BC'')^{\frac{2}{2+\gamma}(2-\beta+\frac{2\beta}{q+1})} \leq 1, \quad (5.138)$$

where  $C' = C'_{(5.136)}(\beta) \geq 1$  and  $C'' = C''_{(5.137)}(\beta) \geq 1$  were already fixed. We recognize (5.138) as the condition on  $\Lambda$  present in Proposition 5.5.

Then, since  $m_* \geq 1$ , we deduce from (5.138) that  $C'\varepsilon_{m_*}^\delta \leq C'\varepsilon_1^\delta \leq C'\Lambda^{-\delta} \leq A$ . Therefore, if the initial datum  $\theta_0$  satisfies

$$\frac{\|\nabla\theta_0\|_{L^2(\mathbb{T}^2)}^4}{\|\theta_0\|_{L^2(\mathbb{T}^2)}^2 \|\Delta\theta_0\|_{L^2(\mathbb{T}^2)}^2} \geq A, \quad (5.139)$$

then (5.136) is satisfied. In order to obey condition (5.135), we additionally assume that  $\theta_0$  satisfies

$$(C'')^{\frac{2+\gamma-2q\delta}{2+\gamma}} \varepsilon_{m_*}^{\frac{2+\gamma-2q\delta}{2}} \leq \frac{\|\theta_0\|_{L^2(\mathbb{T}^2)}}{\|\nabla\theta_0\|_{L^2(\mathbb{T}^2)}} \leq (BC'')^{\frac{2+\gamma-2q\delta}{2+\gamma}} \varepsilon_{m_*}^{\frac{2+\gamma-2q\delta}{2}}, \quad (5.140)$$

where  $C'' = C''_{(5.137)}(\beta) \geq 1$ , and we recall that  $L_{\theta_0} = \|\theta_0\|_{L^2(\mathbb{T}^2)} \|\nabla\theta_0\|_{L^2(\mathbb{T}^2)}^{-1}$ . We remark that the lower bound in (5.140) is exactly a re-statement of the lower bound in (5.137), so this lower bound does not constitute an assumption (this is just part of the definition of  $m_*$ ); instead, it is the upper bound in (5.140) which is the assumption, and it states that this upper bound matches the lower bound up to some a-priori fixed constant (the upper bound in (5.137) does not guarantee this fact). With this extra assumption on  $\theta_0$  imposed, we note that (5.138) and (5.140) imply

$$\begin{aligned} (C'\varepsilon_{m_*}^\delta)^{\beta-2-\frac{2\beta}{q+1}} &\leq (C'\Lambda^{-\delta})(BC'')^{\frac{2}{2+\gamma}(2-\beta+\frac{2\beta}{q+1})} L_{\theta_0}^{-(1+\frac{2q\delta}{2+\gamma-2q\delta})\frac{2}{2+\gamma}(2-\beta+\frac{2\beta}{q+1})} \\ &\leq L_{\theta_0}^{-(1+\frac{2q\delta}{2+\gamma-2q\delta})\frac{2}{2+\gamma}(2-\beta+\frac{2\beta}{q+1})} \leq L_{\theta_0}^{-2}. \end{aligned}$$

The last inequality holds since  $L_{\theta_0} \leq 1/(2\pi) \leq 1$ , and upon unpacking the definitions of  $q, \gamma, \delta$ , as functions of  $\beta$  (see (2.2)–(2.7)), we observe that  $(1 + \frac{2q\delta}{2+\gamma-2q\delta})\frac{2}{2+\gamma}(2 - \beta + \frac{2\beta}{q+1}) \leq 2$ , whenever

$68/67 \leq \beta < 4/3$ . Since the right side of (5.136) is by definition equal to  $L_{\theta_0}^{-2}$ , the above estimate shows that condition (5.136) is indeed satisfied. We recognize (5.139)–(5.140) as being precisely the conditions on  $\theta_0$  present in the statement of Proposition 5.5.

To summarize, we have proven that for  $\Lambda$  sufficiently large to obey (5.138), and all initial data  $\theta_0 \in \dot{H}^2(\mathbb{T})$  which satisfies (5.139)–(5.140), there exists an integer  $m_* \geq 1$  and a nonempty interval of times  $t \in [0, 1]$ , such that the solutions  $\{\theta_m^{(1)}\}_{m \geq m_*}$  and  $\{\theta_m^{(2)}\}_{m \geq m_*}$  constructed in (5.124) satisfy (5.134), i.e.

$$|\|\theta_m^{(1)}(t, \cdot)\|_{L^2(\mathbb{T}^2)} - \|\theta_m^{(2)}(t, \cdot)\|_{L^2(\mathbb{T}^2)}| \geq C(\beta, L_{\theta_0}) \|\theta_0\|_{L^2(\mathbb{T}^2)} > 0.$$

Recall however that the functions  $\theta_m^{(i)}$  defined in (5.124) are not the same as the functions  $\theta_m^{\kappa_m^{(i)}}$ , which are the solutions of the drift-diffusion equation with diffusivity parameter  $\kappa_m^{(i)}$  and full vector field  $\mathbf{b}$  (instead of the truncated one  $\mathbf{b}_m$ ). To estimate this difference we apply (5.101). In the context of the present argument, this estimate is to be applied with  $\theta \mapsto \theta_m^{\kappa_m^{(i)}}$  and  $\theta_M \mapsto \theta_m^{(i)}$ , and  $M \mapsto m$ . We deduce that for all  $m \geq m_*$ ,

$$\|\theta_m^{\kappa_m^{(i)}} - \theta_m^{(i)}\|_{L^\infty((0,1); L^2(\mathbb{T}^2))} \leq C \varepsilon_m^{2\beta(q - \frac{2}{q+1})} \|\theta_0\|_{L^2(\mathbb{T}^2)}, \quad (5.141)$$

for some  $C = C(\beta) > 0$ . Since the exponent  $2\beta(q - \frac{2}{q+1})$  is positive, for  $m$  sufficiently large depending on  $\beta$  and  $\theta_0$ , we can ensure that  $C(\beta) \varepsilon_m^{2\beta(q - \frac{2}{q+1})} \leq \frac{1}{2} C(\beta, L_{\theta_0})$ , and thus for some  $t \in [0, 1]$  and all  $m$  sufficiently large,

$$|\|\theta_m^{\kappa_m^{(1)}}(t, \cdot)\|_{L^2(\mathbb{T}^2)} - \|\theta_m^{\kappa_m^{(2)}}(t, \cdot)\|_{L^2(\mathbb{T}^2)}| \geq \frac{1}{2} C(\beta, L_{\theta_0}) \|\theta_0\|_{L^2(\mathbb{T}^2)} > 0. \quad (5.142)$$

Moreover, estimate (5.90) implies that each of the two sequences  $\{\theta_m^{(1)}\}_{m \geq m_*}$  and  $\{\theta_m^{(2)}\}_{m \geq m_*}$  are Cauchy in  $C^{0,\mu}([0, 1]; L^2(\mathbb{T}^2))$  for some  $\mu > 0$ , since the  $\varepsilon_m$ 's decay super-geometrically as  $m \rightarrow \infty$ . By the triangle inequality and (5.141), the sequences  $\{\theta_m^{\kappa_m^{(1)}}\}_{m \geq m_*}$  and  $\{\theta_m^{\kappa_m^{(2)}}\}_{m \geq m_*}$  are Cauchy in  $L^\infty([0, 1]; L^2(\mathbb{T}^2))$ ; in fact, in light of (5.108) they are Cauchy sequences in  $C^{0,\mu/2}([0, 1]; L^2(\mathbb{T}^2))$ . Clearly, these two sequences have different limits due to (5.142). Moreover, these limit points are weak solutions of the transport equation with velocity field  $\mathbf{b}$ , and belong to  $C^{0,\mu/2}([0, 1]; L^2(\mathbb{T}^2))$ . This concludes the proof of Proposition 5.5.  $\square$

## A. Macroscopic mean drift destroys enhancement

The goal is to formalize the idea, mentioned in the introduction, that a slowly-varying background flow with large amplitude will destroy the enhancement generated by a mean-zero, time-independent microscopic flow with smaller amplitude. The arguments in this appendix are not used anywhere in the paper; we include these for informational purposes only.

To simplify the discussion, we consider only two scales and assume that the macroscopic background flow is constant. We therefore assume that  $\mathbf{b} : \mathbb{T}^d \rightarrow \mathbb{R}^d$  is a periodic, incompressible, mean-zero, time-independent vector field and  $v \in \mathbb{R}^d$  is a vector representing the constant background flow. We let  $\mathbf{m}$  be the stream matrix for  $\mathbf{b}$  and define for  $0 < \varepsilon \ll 1$

$$\mathbf{a}^\varepsilon := I_d + \mathbf{m}(\dot{\varepsilon}).$$

We may then write

$$-\Delta + (\frac{1}{\varepsilon} \mathbf{b}(\frac{x}{\varepsilon}) + \frac{1}{\varepsilon} v) \cdot \nabla = -\nabla \cdot \mathbf{a}^\varepsilon(x) \nabla + \frac{1}{\varepsilon} v \cdot \nabla.$$

We wish to estimate the effective diffusivity matrix, denoted by  $\bar{\mathbf{a}}_v$ , which we obtain by homogenizing the operator

$$\partial_t - \nabla \cdot \mathbf{a}^\varepsilon \nabla + \frac{1}{\varepsilon} v \cdot \nabla. \quad (\text{A.1})$$

We can rewrite this operator, absorbing the constant vector  $v$  into the diffusion matrix, by changing variables in space-time: if  $\theta_\varepsilon(t, x)$  is a solution of

$$\partial_t \theta_\varepsilon - \nabla \cdot \mathbf{a}^\varepsilon(x) \nabla \theta_\varepsilon + \frac{1}{\varepsilon} v \cdot \nabla \theta_\varepsilon = 0,$$

then defining  $T_\varepsilon(t, x) := \theta_\varepsilon(t, x - \frac{1}{\varepsilon} tv)$ , we find that

$$\partial_t T_\varepsilon - \nabla \cdot \mathbf{a}_v^\varepsilon \nabla T_\varepsilon = 0, \quad \text{where} \quad \mathbf{a}_v^\varepsilon(t, x) := \mathbf{a}^\varepsilon(x - \frac{1}{\varepsilon} tv) = I_d + \mathbf{m}(\frac{x}{\varepsilon} - v \frac{t}{\varepsilon^2}) =: \mathbf{a}_v(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon}).$$

We can therefore pose our problem as follows: we are interested in computing the effective diffusivity matrix for the parabolic operator  $\partial_t - \nabla \cdot \mathbf{a}_v^\varepsilon(t, x) \nabla$ , where  $\mathbf{a}_v^\varepsilon(t, x) = \mathbf{a}_v(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon})$  is periodic in  $x$  and *quasiperiodic* in the time variable  $t$ . By classical homogenization theory, this problem homogenizes to  $\partial_t - \nabla \cdot \bar{\mathbf{a}}_v \nabla$  for an effective matrix  $\bar{\mathbf{a}}_v$  which depends on  $v$ ; we can invert the change of variables and see that  $\bar{\mathbf{a}}_v$  is also the effective diffusivity of the original operator we were interested in (A.1).

Let us assume that  $v \in \mathbb{R}^d$  is a good Diophantine direction,<sup>14</sup> i.e., that there exist  $A \in (0, 1)$  and  $\kappa > 0$  with

$$\left| \frac{k}{|k|} \cdot \frac{v}{|v|} \right| \geq A |k|^{-\kappa}, \quad \forall k \in \mathbb{Z}_*^d = \mathbb{Z}^d \setminus \{0\}.$$

Note that the set of directions satisfying this condition has full Lebesgue measure. We can write the equation for the correctors  $\chi_e$  associated to the homogenization problem as

$$\partial_t \chi_e - \nabla \cdot \mathbf{a}_v \nabla \chi_e = \mathbf{b}_v \cdot e \quad \text{in } \mathbb{R} \times \mathbb{R}^d, \quad \langle \chi_e \rangle = 0,$$

where  $\mathbf{b}_v(t, x) = \nabla \cdot \mathbf{a}_v(t, x) = \mathbf{b}(x - tv)$ . Here, and in what follows in this appendix, the brackets  $\langle \cdot \rangle$  denote the mean of a quasiperiodic function of time, which is  $\mathbb{T}^d$ -periodic in space, i.e., for such  $f$ ,

$$\langle f \rangle = \lim_{T \rightarrow \infty} \int_{-T}^T \int_{\mathbb{T}^d} f(t, x) dx dt.$$

The enhancement of diffusivity is related to the correctors' gradient field  $\nabla \chi_e$  by the formula

$$e \cdot \bar{\mathbf{a}}_v e - |e|^2 = \langle |\nabla \chi_e|^2 \rangle. \quad (\text{A.2})$$

To obtain (A.2), we use the equation for  $\chi_e = \chi_e(t, x)$  to get

$$0 = -\langle \frac{1}{2} \partial_t \chi_e^2 \rangle = \langle \nabla \chi_e \cdot \mathbf{a}_v (e + \nabla \chi_e) \rangle.$$

Therefore, since  $\langle \nabla \chi_e \rangle = 0$ , and the symmetric part of  $\mathbf{a}_v$  is  $I_d$ , we obtain from the above identity that

$$e \cdot \bar{\mathbf{a}}_v e = e \cdot \langle \mathbf{a}_v (e + \nabla \chi_e) \rangle = \langle (e + \nabla \chi_e) \cdot \mathbf{a}_v (e + \nabla \chi_e) \rangle = \langle |e + \nabla \chi_e|^2 \rangle = |e|^2 + \langle |\nabla \chi_e|^2 \rangle.$$

The identity (A.2) says that the difference between the effective diffusivity  $\bar{\mathbf{a}}_v$  and the molecular diffusivity matrix  $I_d$  is proportional to the size of the correctors. Next, we show that  $\langle |\nabla \chi_e|^2 \rangle$  is small when  $|v| \gg 1$ . For this purpose, we use the identity

$$\langle |\nabla \chi_e|^2 \rangle = \langle (\mathbf{b}_v \cdot e) \chi_e \rangle$$

---

<sup>14</sup>For example, let  $v = (1, \sqrt{2})$  for  $d = 2$ . Then, for any  $k \in \mathbb{Z}_*^2$ ,  $\left| \frac{k}{|k|} \cdot \frac{v}{|v|} \right| = \frac{|k_1 + k_2 \sqrt{2}|}{\sqrt{3}|k|} \geq \frac{1}{3\sqrt{3}|k|^2}$ .

and estimate the right side.

Assume  $f$  and  $g$  are quasiperiodic in  $t$  and  $\mathbb{T}^d$ -periodic in  $x$ . Then, denoting  $\hat{f}_k(t)$  and  $\hat{g}_k(t)$  the Fourier-series coefficients of  $f$  and  $g$  with respect to  $x$  only, we have by Plancherel that

$$\langle f g \rangle = \lim_{T \rightarrow \infty} \int_{-T}^T \sum_{k \in \mathbb{Z}_*^d} \hat{f}_k(t) \overline{\hat{g}_k(t)} dt.$$

We wish to apply this identity to  $f = \chi_e$  and  $g = e \cdot \mathbf{b}_v$ , i.e., we wish to compute  $\langle (\mathbf{b}_v \cdot e) \chi_e \rangle$ . Since translation in physical space is modulation in Fourier space, we know that

$$\overline{\hat{g}_k(t)} = \widehat{(e \cdot \mathbf{b})}_k e^{2\pi i k \cdot v t} = \frac{d}{dt} \left( \frac{-i}{2\pi k \cdot v} \widehat{(e \cdot \mathbf{b})}_k e^{2\pi i k \cdot v t} \right) = \frac{-i}{2\pi k \cdot v} \frac{d}{dt} \overline{\hat{g}_k(t)}.$$

Now, quasiperiodic functions with mean zero gradients are sublinear at infinity, and so we have via integration by parts in time,

$$\langle f g \rangle = \lim_{T \rightarrow \infty} \int_{-T}^T \sum_{k \in \mathbb{Z}_*^d} \hat{f}_k(t) \overline{\hat{g}_k(t)} dt = \lim_{T \rightarrow \infty} \int_{-T}^T \sum_{k \in \mathbb{Z}_*^d} \frac{i}{2\pi k \cdot v} \overline{\hat{g}_k(t)} \frac{d}{dt} \hat{f}_k(t) dt. \quad (\text{A.3})$$

Using the equation satisfied by  $f = \chi_e$ , we have that

$$\frac{d}{dt} \hat{f}_k(t) = 2\pi i k \cdot (\widehat{\mathbf{a}_v \cdot \nabla \chi_e})_k(t) + \hat{g}_k(t).$$

Since  $k \cdot v$  is odd in  $k$ , the second term in the above display does not contribute to the expression in (A.3). We deduce finally that

$$\langle (\mathbf{b}_v \cdot e) \chi_e \rangle = \frac{-1}{|v|} \lim_{T \rightarrow \infty} \int_{-T}^T \sum_{k \in \mathbb{Z}_*^d} \frac{1}{|k| \cdot \frac{v}{|v|}} \widehat{(e \cdot \mathbf{b})}_k e^{2\pi i k \cdot v t} \frac{k}{|k|} \cdot (\widehat{\mathbf{a}_v \cdot \nabla \chi_e})_k(t) dt. \quad (\text{A.4})$$

Now, since  $v$  is a good Diophantine direction, we have the bound

$$\left| \frac{1}{\frac{k}{|k|} \cdot \frac{v}{|v|}} \right| \leq \frac{|k|^\kappa}{A}, \quad \text{for all } k \in \mathbb{Z}_*^d.$$

With this information, we return to (A.4), use Plancherel and Cauchy-Schwartz to deduce that

$$\begin{aligned} \langle |\nabla \chi_e|^2 \rangle &= |\langle (\mathbf{b}_v \cdot e) \chi_e \rangle| \leq \frac{1 + \|\mathbf{m}\|_{L^\infty(\mathbb{T}^d)}}{A|v|} \langle |\nabla \chi_e|^2 \rangle^{1/2} \left( \sum_{k \in \mathbb{Z}_*^d} |k|^{2\kappa} |\widehat{(e \cdot \mathbf{b})}_k|^2 \right)^{1/2} \\ &\leq \frac{1 + \|\mathbf{m}\|_{L^\infty(\mathbb{T}^d)}}{A|v|} \langle |\nabla \chi_e|^2 \rangle^{1/2} \|e \cdot \mathbf{b}\|_{\dot{H}^\kappa(\mathbb{T}^d)}. \end{aligned}$$

As a consequence,

$$\langle |\nabla \chi_e|^2 \rangle \leq \frac{(1 + \|\mathbf{m}\|_{L^\infty(\mathbb{T}^d)})^2 \|e \cdot \mathbf{b}\|_{\dot{H}^\kappa(\mathbb{T}^d)}^2}{A^2 |v|^2}$$

which becomes arbitrarily small for  $|v| \gg 1$ . Returning to (A.2), we have thus shown that a mean drift of “generic” direction destroys the enhancement of diffusion when  $|v| \gg 1$ .

## B. Faá di Bruno formula and its consequences

In order to show that the stream function  $\phi(t, x)$  constructed in Section 2 has  $C^{1,\beta}$  regularity as stated in Proposition 2.2, we require explicit estimates on the derivatives of the solutions  $X$  of the flow  $\partial_t X = \mathbf{f}(t, X)$  in terms of those of  $\mathbf{f}$ . A qualitative version of such an estimate (for instance, the statement that  $\mathbf{f} \in C^k \implies X \in C^k$ ) is of course quite classical and can be found in most introductory textbooks on ODE theory. The difference here is that we need an explicit estimate which, while it must be known, is of a form we could not find written in the literature. Like the qualitative arguments, the proof boils down to differentiating the equation many times. The only difficulty is a bookkeeping one: we must keep track of all the terms arising out of repeatedly applying the chain rule; that is, we need to use the Faá di Bruno formula.

### B.1 Useful lemmas

Recall from (1.21) that for any integer  $n \geq 0$ ,  $C_x^n$  smooth function  $f$ , and  $R > 0$ , we denote

$$[\![f]\!]_{n,R} = \frac{(n+1)^2}{n!R^n} \sup_{|\alpha|=n} \|\partial^\alpha f\|_{L_x^\infty}. \quad (\text{B.1})$$

In the above definition, the shift factor  $(n+1)^2$  could be replaced by  $(n+1)^r$  for any  $r > 1$ , at the cost of introducing  $r$ -dependence on our constants. We shall frequently use the identity

$$\sum_{|\beta|=k, \beta \leq \alpha} \binom{\alpha}{\beta} = \binom{n}{k} \quad (\text{B.2})$$

for  $|\alpha| = n$  and  $0 \leq k \leq n$ . By combining this identity with the Leibniz rule, we obtain:

**Lemma B.1 (Product estimate).** *For  $f, g \in C_x^n$  we have the bound*

$$[\![fg]\!]_{n,R} \leq 4 \left( \max_{0 \leq j \leq n} [\![f]\!]_{j,R} \right) \left( \max_{0 \leq j \leq n} [\![g]\!]_{j,R} \right). \quad (\text{B.3})$$

*Proof of Lemma B.1.* The Leibniz rule and (B.2) give

$$\begin{aligned} [\![fg]\!]_{n,R} &\leq \frac{(n+1)^2}{n!R^n} \sum_{k=0}^n \sup_{|\alpha|=n} \sum_{|\beta|=k, \beta \leq \alpha} \binom{\alpha}{\beta} \|\partial^\beta f\|_{L_x^\infty} \|\partial^{\alpha-\beta} g\|_{L_x^\infty} \\ &\leq \frac{(n+1)^2}{n!R^n} \sum_{k=0}^n \binom{n}{k} [\![f]\!]_{k,R} \frac{k!R^k}{(k+1)^2} [\![g]\!]_{n-k,R} \frac{(n-k)!R^{n-k}}{(n-k+1)^2} \\ &\leq \left( \max_{0 \leq j \leq n} [\![f]\!]_{j,R} \right) \left( \max_{0 \leq j \leq n} [\![g]\!]_{j,R} \right) \sum_{k=0}^n \frac{(n+1)^2}{(k+1)^2(n-k+1)^2}. \end{aligned}$$

The proof now follows since  $\sum_{k=0}^n (n+1)^2(k+1)^{-2}(n-k+1)^{-2} \leq 4$  for all  $n \geq 0$ .  $\square$

The workhorse of this Appendix is a multivariable version of the Faá di Bruno formula, with requires some additional notation. We denote by  $\mathbb{N}_0$  the set of all integers strictly larger than  $-1$ , and by  $\mathbb{N}_0^d$  the set of all multi-indices  $\alpha = (\alpha_1, \dots, \alpha_d)$  with  $\alpha_j \in \mathbb{N}_0$ . For a multi-index  $\alpha$ , we write  $|\alpha| = \alpha_1 + \dots + \alpha_d$ ,  $\alpha! = (\alpha_1!) \cdot \dots \cdot (\alpha_d!)$ ,  $\partial^\alpha = \partial_{x_1}^{\alpha_1} \dots \partial_{x_d}^{\alpha_d}$ , and  $\mathbf{y}^\alpha = (y_1^{\alpha_1}) \cdot \dots \cdot (y_d^{\alpha_d})$ , where

$\mathbf{y} \in \mathbb{R}^d$  is a point. The following notation shall be needed below. Let  $n \geq 1$ ,  $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}_0^d$  be such that  $|\boldsymbol{\beta}| = n$ , and  $1 \leq |\boldsymbol{\alpha}| \leq n$ . For  $1 \leq s \leq n$  define the set<sup>15</sup>

$$p_s(\boldsymbol{\beta}, \boldsymbol{\alpha}) = \left\{ (\mathbf{k}_1, \dots, \mathbf{k}_s; \boldsymbol{\ell}_1, \dots, \boldsymbol{\ell}_s) \in (\mathbb{N}_0^d \times \dots \mathbb{N}_0^d; \mathbb{N}_0^d \times \dots \mathbb{N}_0^d) : \right. \\ \left. 0 < |\mathbf{k}_j|, 0 < \boldsymbol{\ell}_1 < \dots < \boldsymbol{\ell}_s, \sum_{j=1}^s \mathbf{k}_j = \boldsymbol{\alpha}, \sum_{i=1}^s |\mathbf{k}_j| \boldsymbol{\ell}_j = \boldsymbol{\beta} \right\}. \quad (\text{B.4})$$

With this notation in hand, we recall [CS96, Theorem 2.1].

**Proposition B.2 (Multivariate Faà di Bruno Formula).** *Let  $h: \mathbb{R}^d \rightarrow \mathbb{R}$  be  $C^\infty$  in a neighborhood of  $\mathbf{y}_0 := \mathbf{g}(\mathbf{x}_0)$  and  $\mathbf{g}: \mathbb{R}^d \rightarrow \mathbb{R}^d$  be  $C^\infty$  in a neighborhood of  $\mathbf{x}_0$ . Denote their composition by  $f := h \circ \mathbf{g}$ . Then, for every multiindex  $\boldsymbol{\beta}$  with  $n := |\boldsymbol{\beta}| \geq 1$ ,*

$$(\partial^{\boldsymbol{\beta}} f)(\mathbf{x}_0) = \boldsymbol{\beta}! \sum_{1 \leq |\boldsymbol{\alpha}| \leq n} (\partial^{\boldsymbol{\alpha}} h)(\mathbf{g}(\mathbf{x}_0)) \sum_{s=1}^n \sum_{\mathbf{k}_j \in p_s(\boldsymbol{\beta}, \boldsymbol{\alpha})} \prod_{j=1}^s \frac{((\partial^{\boldsymbol{\ell}_j} \mathbf{g})(\mathbf{x}_0))^{\mathbf{k}_j}}{(\mathbf{k}_j!) (\boldsymbol{\ell}_j!)^{|\mathbf{k}_j|}}.$$

Here we adopt the convention that  $0^0 := 1$ .

Proposition B.2 expands the higher order chain rule into a complicated expression. In applications, we need to have tools which allow us to contract such complicated expressions into simple ones. The first re-summation lemma that we use in the paper is similar to [CVW15, Lemma 3.2]:

**Lemma B.3.** *Fix the dimension  $d \geq 1$ . With the notation of Proposition B.2, we have*

$$\boldsymbol{\beta}! \sum_{1 \leq |\boldsymbol{\alpha}| \leq n} (-d)^{-|\boldsymbol{\alpha}|} |\boldsymbol{\alpha}|! \sum_{s=1}^n \sum_{\mathbf{k}_j \in p_s(\boldsymbol{\beta}, \boldsymbol{\alpha})} \prod_{j=1}^s \frac{(|\boldsymbol{\ell}_j|! (\boldsymbol{\ell}_j!)^{1/2})^{|\mathbf{k}_j|}}{(\mathbf{k}_j!) (\boldsymbol{\ell}_j!)^{|\mathbf{k}_j|}} = 2(n+1)! \binom{1/2}{n+1}.$$

*Proof of Lemma B.3.* Define the following functions

$$h(\mathbf{y}) = \bar{h}(y_1 + \dots + y_d), \quad \bar{h}(z) = \frac{1}{1 - \frac{z}{d}},$$

$$g_1(\mathbf{x}) = \dots = g_d(\mathbf{x}) = \bar{g}(x_1 + \dots + x_d), \quad \bar{g}(z) = 1 - \sqrt{1 - z},$$

$$f(\mathbf{x}) = h(g_1(\mathbf{x}), \dots, g_d(\mathbf{x})) = \bar{f}(x_1 + \dots + x_d), \quad \bar{f}(z) = \bar{h}(d\bar{g}(z)) = \frac{1}{\sqrt{1 - z}}.$$

which are real-analytic functions in the neighborhood of  $\mathbf{x} = 0$ . For any multi-index  $\boldsymbol{\alpha} \in \mathbb{N}_0^d$  we have that  $(\partial^{\boldsymbol{\alpha}} f)(0) = (\partial^{|\boldsymbol{\alpha}|} \bar{f})(0)$  and similarly for the functions  $h, g_1, \dots, g_d$ . Moreover, we note the following identities<sup>16</sup>

$$(\partial^n \bar{h})(0) = d^{-n} n! \quad (\text{B.5a})$$

$$(\partial^n \bar{g})(0) = (-1)^{n-1} \binom{1/2}{n} n! \geq 0 \quad (\text{B.5b})$$

$$(\partial^n \bar{f})(0) = 2(-1)^n \binom{1/2}{n+1} (n+1)! \geq 0 \quad (\text{B.5c})$$

<sup>15</sup>Here  $\boldsymbol{\ell}_1 < \boldsymbol{\ell}_2$  if either  $|\boldsymbol{\ell}_1| < |\boldsymbol{\ell}_2|$ , or  $|\boldsymbol{\ell}_1| = |\boldsymbol{\ell}_2|$  and there exists  $k \in \{1, \dots, d\}$  such that  $\boldsymbol{\ell}_{1k'} = \boldsymbol{\ell}_{2k'}$  for  $k' < k$ , and  $\boldsymbol{\ell}_{1k} < \boldsymbol{\ell}_{2k}$ .

<sup>16</sup>Recall that  $\binom{1/2}{n} n! = (1/2)(1/2 - 1) \dots (1/2 - n + 1)$  for  $n \geq 1$ , and  $\binom{1/2}{0} := -1$ .

which are valid for  $n \geq 1$ . Let  $\beta$  be any multi-index of length  $n$ . We apply Proposition B.2 to the function  $f(x)$  defined above, and using (B.5a)–(B.5b) deduce that

$$\begin{aligned} (-1)^{-n}(\partial^\beta f)(0) &= (-1)^{-n}\beta! \sum_{1 \leq |\alpha| \leq n} d^{-|\alpha|} |\alpha|! \sum_{s=1}^n \sum_{p_s(\beta, \alpha)} \prod_{j=1}^s \frac{\left((-1)^{|\ell_j|-1} \binom{1/2}{|\ell_j|} |\ell_j|!\right)^{|\mathbf{k}_j|}}{(\mathbf{k}_j!)(\ell_j!)^{|\mathbf{k}_j|}} \\ &= \beta! \sum_{1 \leq |\alpha| \leq n} (-d)^{-|\alpha|} |\alpha|! \sum_{s=1}^n \sum_{p_s(\beta, \alpha)} \prod_{j=1}^s \frac{\left(\binom{1/2}{|\ell_j|} |\ell_j|!\right)^{|\mathbf{k}_j|}}{(\mathbf{k}_j!)(\ell_j!)^{|\mathbf{k}_j|}} \end{aligned} \quad (\text{B.6})$$

which is the expression that we wish to estimate. Here we have used that  $\sum_j |\mathbf{k}_j| = |\alpha|$  and  $\sum_j |\mathbf{k}_j| |\ell_j| = |\beta| = n$ . On the other hand, we know that from (B.5c) that

$$(-1)^{-n}(\partial^\beta f)(0) = (-1)^{-n}(\partial^n \bar{f})(0) = 2 \binom{1/2}{n+1} (n+1)! \quad (\text{B.7})$$

Equating (B.6) and (B.7) concludes the proof of the lemma.  $\square$

Another re-summation lemma that we use is a multi-D version of [KP02, Lemma 1.4.1]:

**Lemma B.4.** *Let  $R > 0$  and  $d \geq 1$ . With the notation of Proposition B.2, we have that*

$$\beta! \sum_{1 \leq |\alpha| \leq n} R^{|\alpha|} |\alpha|! \sum_{s=1}^n \sum_{p_s(\beta, \alpha)} \prod_{j=1}^s \frac{(|\ell_j|!)^{|\mathbf{k}_j|}}{(\mathbf{k}_j!)(\ell_j!)^{|\mathbf{k}_j|}} = dR(1+dR)^{n-1}n!$$

*Proof of Lemma B.4.* Similarly to the proof of Lemma B.3, we define the functions

$$\begin{aligned} h(\mathbf{y}) &= \bar{h}(y_1 + \dots + y_d), & \bar{h}(z) &= \frac{1}{1-dR(z-1)}, \\ g_i(\mathbf{x}) &= \bar{g}(x_1 + \dots + x_d), & \bar{g}(z) &= \frac{1}{d(1-z)}, \\ f(\mathbf{x}) &= h(g_1(\mathbf{x}), \dots, g_d(\mathbf{x})) = \bar{f}(x_1 + \dots + x_d), & \bar{f}(z) &= \bar{h}(d\bar{g}(z)) = \frac{1-z}{1-(dR+1)z}, \end{aligned}$$

so that for any  $\beta \in \mathbb{N}_0^2$  with  $|\beta| = n$ , we have

$$\begin{aligned} (\partial^\beta h)(1/d, \dots, 1/d) &= (\partial^n \bar{h})(1) = (dR)^n n!, \\ (\partial^\beta g_i)(0, \dots, 0) &= (\partial^n \bar{g}_i)(0) = \frac{n!}{d}, \\ (\partial^\beta f)(0, \dots, 0) &= (\partial^n \bar{f})(0) = dR(1+dR)^{n-1}n!. \end{aligned}$$

Using the above identities and Proposition B.2, we obtain that

$$\begin{aligned} dR(1+dR)^{n-1}n! &= (\partial^\beta f)(0, \dots, 0) \\ &= \beta! \sum_{1 \leq |\alpha| \leq n} (dR)^{|\alpha|} |\alpha|! \sum_{s=1}^n \sum_{p_s(\beta, \alpha)} \prod_{j=1}^s \frac{\left(\frac{1}{d} |\ell_j|!\right)^{|\mathbf{k}_j|}}{(\mathbf{k}_j!)(\ell_j!)^{|\mathbf{k}_j|}} \\ &= \beta! \sum_{1 \leq |\alpha| \leq n} R^{|\alpha|} |\alpha|! \sum_{s=1}^n \sum_{p_s(\beta, \alpha)} \prod_{j=1}^s \frac{(|\ell_j|!)^{|\mathbf{k}_j|}}{(\mathbf{k}_j!)(\ell_j!)^{|\mathbf{k}_j|}} \end{aligned}$$

which concludes the proof.  $\square$

**Remark B.5.** Due to the presence of the shift factor  $(n+1)^2$  in (B.1), in order for Lemmas B.3 and B.4 to be useful, we shall need the following inequality. Assume that  $|\beta| = n$ ,  $1 \leq |\alpha| \leq n$ ,  $1 \leq s \leq n$ , and  $(\mathbf{k}_1, \dots, \mathbf{k}_s; \ell_1, \dots, \ell_s) \in p_s(\beta, \alpha)$ , as defined in (B.4). Then, we have that

$$\frac{(n+1)^2}{(|\alpha|+1)^2} \prod_{j=1}^s \frac{1}{(|\ell_j|+1)^{2|\mathbf{k}_j|}} \leq 1. \quad (\text{B.8})$$

In order to prove (B.8), we note that by the definition of the partition set  $p_s(\beta, \alpha)$  in (B.4), we have  $|\ell_s| \geq |\ell_j| \geq 1$  for all  $1 \leq j \leq s$ . As such, we obtain that  $n = |\beta| = \sum_{j=1}^s |\mathbf{k}_j| |\ell_j| \leq |\ell_s| \sum_{j=1}^s |\mathbf{k}_j| = |\ell_s| |\alpha|$ . Thus, the lower bound  $|\ell_s| \geq n/|\alpha|$  emerges, and since  $|\mathbf{k}_s| \geq 1$ , we have

$$\begin{aligned} \frac{(n+1)^2}{(|\alpha|+1)^2} \prod_{j=1}^s \frac{1}{(|\ell_j|+1)^{2|\mathbf{k}_j|}} &\leq \frac{(n+1)^2}{(|\alpha|+1)^2} \frac{1}{(|\ell_s|+1)^{2|\mathbf{k}_s|}} \\ &\leq \frac{(n+1)^2}{(|\alpha|+1)^2} \frac{1}{\left(\frac{n}{|\alpha|}+1\right)^2} = \frac{|\alpha|^2}{(|\alpha|+1)^2} \frac{(n+1)^2}{(n+|\alpha|)^2} \leq 1. \end{aligned}$$

In the last inequality we have used that  $|\alpha| \geq 1$ . This proves (B.8).

The following estimate is standard for  $m = \infty$ , as it give the radius of analyticity of the composition of two real-analytic functions. We recall it here for the sake of completeness, and note trivially that it holds for finite  $m$ .

**Proposition B.6 (Composition estimate).** *Let  $h \in L^\infty(C^m(\mathbb{R}^d))$ ,  $\mathbf{g} \in L^\infty(C^m(\mathbb{R}^d))^d$ , and assume that there exist positive constants  $C_h, C_g, R_h, R_g \in (0, \infty)$  such that*

$$\|h\|_{n, R_h} \leq C_h \quad \text{and} \quad \|\mathbf{g}\|_{n, R_g} \leq C_g$$

for all  $0 \leq n \leq m$  and respectively  $1 \leq n \leq m$ . Then, for every  $0 \leq n \leq m$ , we have

$$\|h \circ \mathbf{g}\|_{n, R} \leq C_h \quad \text{where} \quad R = R_g(1 + dC_g R_h). \quad (\text{B.9})$$

*Proof of Proposition B.6.* Let  $f(\mathbf{x}) = h(\mathbf{g}(\mathbf{x})) \in L^\infty(C^m(\mathbb{R}^d))$ . For  $n = 0$  the statement (B.9) holds trivially. Let  $n \geq 1$  and fix  $\beta \in \mathbb{N}_0^d$  with  $|\beta| = n$ . From Proposition B.2, the assumed bounds on  $h$  and  $\mathbf{g}$ , the re-summation formula in Lemma B.4, and the bound (B.8), we deduce that

$$\begin{aligned} \|\partial^\beta f\|_{L^\infty} &\leq \beta! \sum_{1 \leq |\alpha| \leq n} C_h R_h^{|\alpha|} \frac{|\alpha|!}{(|\alpha|+1)^2} \sum_{s=1}^n \sum_{p_s(\beta, \alpha)} \prod_{j=1}^s \frac{\left(C_g R_g^{|\ell_j|} \frac{|\ell_j|!}{(|\ell_j|+1)^2}\right)^{|\mathbf{k}_j|}}{(\mathbf{k}_j!) (\ell_j!)^{|\mathbf{k}_j|}} \\ &\leq C_h R_g^n \frac{\beta!}{(|\beta|+1)^2} \sum_{1 \leq |\alpha| \leq n} (C_g R_h)^{|\alpha|} |\alpha|! \sum_{s=1}^n \sum_{p_s(\beta, \alpha)} \prod_{j=1}^s \frac{(|\ell_j|!)^{|\mathbf{k}_j|}}{(\mathbf{k}_j!) (\ell_j!)^{|\mathbf{k}_j|}} \\ &= C_h R_g^n (dC_g R_h) (1 + dC_g R_h)^{n-1} \frac{n!}{(n+1)^2}. \end{aligned} \quad (\text{B.10})$$

Re-arranging the right side of the above, and recalling (B.1), finishes the proof.  $\square$

## B.2 Application to the transport equation

For given smooth functions  $\mathbf{f}, \mathbf{g}: \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  we first consider the solution  $Y$  of

$$\begin{cases} (\partial_t + \mathbf{f} \cdot \nabla) Y = \mathbf{g} & \text{in } \mathbb{R} \times \mathbb{R}^d, \\ Y(0, \cdot) = 0 & \text{on } \mathbb{R}^d. \end{cases} \quad (\text{B.11})$$

**Lemma B.7.** Assume there exist  $C_{\mathbf{f}}, R_{\mathbf{f}}, C_{\mathbf{g}}, R_{\mathbf{g}} > 0$  with  $R_{\mathbf{g}} \geq R_{\mathbf{f}}$  and  $N \in \mathbb{N}$  such that

$$\max_{1 \leq n \leq N} \sup_{t \in \mathbb{R}} \|\mathbf{f}(t, \cdot)\|_{n, R_{\mathbf{f}}} \leq C_{\mathbf{f}}, \quad \text{and} \quad \max_{1 \leq n \leq N} \sup_{t \in \mathbb{R}} \|\mathbf{g}(t, \cdot)\|_{n, R_{\mathbf{g}}} \leq C_{\mathbf{g}}, \quad (\text{B.12})$$

Then, the solution  $Y$  of the transport equation (B.11) satisfies

$$\max_{1 \leq n \leq N} \sup_{t \in [-T, T]} \frac{1}{|t|} \|Y(t, \cdot)\|_{n, R_Y(t)} \leq 8dC_{\mathbf{g}}, \quad (\text{B.13})$$

where

$$R_Y(t) := R_{\mathbf{g}} + 4|t|dC_{\mathbf{f}}R_{\mathbf{f}}^2 \leq R_{\mathbf{g}}(1 + 4|t|dC_{\mathbf{f}}R_{\mathbf{f}}) \quad \text{and} \quad T := \frac{1}{4dC_{\mathbf{f}}R_{\mathbf{f}}}. \quad (\text{B.14})$$

*Proof of Lemma B.7.* For  $n = 1$ , upon differentiating (B.11) we obtain

$$(\partial_t + \mathbf{f} \cdot \nabla) \nabla Y = \nabla \mathbf{g} + \nabla \mathbf{f} \cdot \nabla Y, \quad \nabla Y(0, \cdot) = 0.$$

Integrating this expression and appealing to (B.12), we obtain

$$\|Y(t, \cdot)\|_{1, R_Y} \leq 4R_Y^{-1} \|\nabla Y(t, \cdot)\|_{L_x^\infty} \leq \frac{dC_{\mathbf{g}}R_{\mathbf{g}}|t|e^{\frac{1}{4}dTR_{\mathbf{f}}C_{\mathbf{f}}}}{R_Y} \leq 2dC_{\mathbf{g}}$$

so that (B.13) holds when  $n = 1$ .

We prove (B.13) inductively on  $n$ , and without loss of generality, we only prove it for  $t \in (0, T]$ . Let  $|\alpha| = n$ . Applying  $\partial^\alpha$  to both sides of (B.11) yields:

$$\begin{cases} (\partial_t + \mathbf{f} \cdot \nabla) \partial^\alpha Y = \partial^\alpha \mathbf{g} + \sum_{\beta < \alpha} \binom{\alpha}{\beta} \partial^{\alpha-\beta} \mathbf{f} \cdot \nabla \partial^\beta Y & \text{in } \mathbb{R} \times \mathbb{R}^d, \\ \partial^\alpha Y(0, \cdot) = 0 & \text{on } \mathbb{R}^d. \end{cases} \quad (\text{B.15})$$

Integrating (B.15), using assumption (B.12) and the inductive assumption (B.13), we get

$$\begin{aligned} \sup_{|\alpha|=n} \|\partial^\alpha Y(t, \cdot)\|_{L_x^\infty} &\leq t \sup_{|\alpha|=n} \|\partial^\alpha \mathbf{g}\|_{L_{t,x}^\infty} + C_{\mathbf{f}} \sup_{|\alpha|=n} \sum_{|\beta|=n-1, \beta < \alpha} \sum_{\ell=1}^d \binom{\alpha}{\beta} \frac{R_{\mathbf{f}}}{2^2} \int_0^t \|\partial^{\beta+e_\ell} Y(s, \cdot)\|_{L_x^\infty} ds \\ &\quad + 16d^2 C_{\mathbf{f}} C_{\mathbf{g}} \sum_{k=0}^{n-2} \sup_{|\alpha|=n} \sum_{|\beta|=k, \beta < \alpha} \binom{\alpha}{\beta} \frac{R_{\mathbf{f}}^{n-k}(n-k)!}{(n-k+1)^2} \frac{(k+1)!}{(k+2)^2} \int_0^t s R_Y(s)^{k+1} ds \\ &\leq t C_{\mathbf{g}} R_{\mathbf{g}}^n \frac{n!}{(n+1)^2} + \frac{dC_{\mathbf{f}}R_{\mathbf{f}}n}{4} \int_0^t \sup_{|\alpha|=n} \|\partial^\alpha Y(s, \cdot)\|_{L_x^\infty} ds \\ &\quad + 16d^2 C_{\mathbf{f}} C_{\mathbf{g}} \frac{n!}{(n+1)^2} \sum_{k=0}^{n-2} \binom{n}{k} \frac{R_{\mathbf{f}}^{n-k}(n+1)^2(k+1)}{(n-k+1)^2(k+2)^2} \int_0^t s R_Y(s)^{k+1} ds. \end{aligned}$$

In the second inequality above we have appealed to identity (B.2). At this point we note that due to the definition of  $R_Y(s)$  in (B.14), we have

$$\int_0^t s R_Y(s)^j ds \leq \frac{t R_Y(t)^{j+1}}{4d(j+1)C_{\mathbf{f}}R_{\mathbf{f}}^2}. \quad (\text{B.16})$$

It follows from the above two estimates, the definition (B.1), and the fact that by (B.14) we have  $R_Y(t) \geq R_{\mathbf{g}} \geq R_{\mathbf{f}}$ , we obtain

$$\begin{aligned} \frac{\llbracket Y(\cdot, t) \rrbracket_{n, R_Y(t)}}{t} &\leq \frac{C_{\mathbf{g}} R_{\mathbf{g}}^n}{R_Y(t)^n} + \frac{dC_{\mathbf{f}} R_{\mathbf{f}} n}{4t R_Y(t)^n} \int_0^t s R_Y(s)^n ds \sup_{s \in [0, t]} \frac{\llbracket Y(\cdot, s) \rrbracket_{n, R_Y(s)}}{s} \\ &\quad + \frac{16d^2 C_{\mathbf{f}} C_{\mathbf{g}}}{t R_Y(t)^n} \sum_{k=0}^{n-2} \frac{R_{\mathbf{f}}^{n-k} (n+1)^2 (k+1)}{(n-k+1)^2 (k+2)^2} \int_0^t s R_Y(s)^{k+1} ds \\ &\leq C_{\mathbf{g}} + \frac{1}{16} \sup_{s \in [0, t]} \frac{\llbracket Y(\cdot, s) \rrbracket_{n, R_Y(s)}}{s} + 4C_{\mathbf{g}} d \sum_{k=0}^{n-2} \frac{(n+1)^2}{(n-k+1)^2 (k+2)^2}. \end{aligned}$$

Using the estimate  $\sum_{k=0}^{n-2} (n+1)^2 (k+1)^{-2} (n-k+1)^{-2} \leq 3/2$ , and taking the supremum over  $t \in [0, T]$ , we obtain that

$$\sup_{t \in (0, T]} t^{-1} \llbracket Y(\cdot, t) \rrbracket_{n, R_Y(t)} \leq \frac{16}{15} (C_{\mathbf{g}} + 6dC_{\mathbf{g}}) \leq 8dC_{\mathbf{g}}.$$

The above estimate shows that (B.13) holds at level  $n$ , closing the inductive step.  $\square$

**Corollary B.8.** *Under the assumptions of Lemma B.7, for all  $|t| \leq T$  and  $0 \leq n \leq N-1$ , we have that*

$$\llbracket \nabla Y(\cdot, t) \rrbracket_{n, R_{\nabla Y}(t)} \leq \frac{4dC_{\mathbf{g}} R_{\mathbf{g}}}{C_{\mathbf{f}} R_{\mathbf{f}}} \quad (\text{B.17})$$

where

$$R_{\nabla Y}(t) = R_{\mathbf{g}} (1 + 4|t|dC_{\mathbf{f}} R_{\mathbf{f}})^2 = R_Y(t) (1 + 4|t|dC_{\mathbf{f}} R_{\mathbf{f}}). \quad (\text{B.18})$$

*Proof of Corollary B.8.* Using the definition (B.1), the bound (B.13), and the definitions of  $R_Y(t)$  in (B.14), and  $R_{\nabla Y}(t)$  in (B.18), we obtain

$$\begin{aligned} \llbracket \nabla Y(\cdot, t) \rrbracket_{n, R_{\nabla Y}(t)} &\leq \frac{d(n+1)^2}{n! R_{\nabla Y}(t)^n} \frac{(n+1)! R_Y(t)^{n+1}}{(n+2)^2} \llbracket Y(\cdot, t) \rrbracket_{n+1, R_Y(t)} \\ &\leq \frac{dn R_Y(t)^{n+1}}{R_{\nabla Y}(t)^n} 8tdC_{\mathbf{g}} \leq \frac{16d^2 t n C_{\mathbf{g}} R_{\mathbf{g}}}{(1 + 4tdC_{\mathbf{f}} R_{\mathbf{f}})^n} \leq \frac{4dC_{\mathbf{g}} R_{\mathbf{g}}}{C_{\mathbf{f}} R_{\mathbf{f}}}. \end{aligned}$$

In the last inequality we have used that  $na \leq (1+a)^n$  for any  $a \geq 0$ .  $\square$

**Lemma B.9.** *If in addition to (B.12) we assume that*

$$\llbracket \partial_t^m \mathbf{f} \rrbracket_{n, R_{\mathbf{f}}} \leq C_{\mathbf{f}} Q_{\mathbf{f}}^m m!, \quad \text{and} \quad \llbracket \partial_t^m \mathbf{g} \rrbracket_{n, R_{\mathbf{g}}} \leq C_{\mathbf{g}} Q_{\mathbf{g}}^m m!, \quad (\text{B.19})$$

for all  $0 \leq n \leq N$  and  $0 \leq m \leq M$ , for some constants  $Q_{\mathbf{g}}, Q_{\mathbf{f}} > 0$ , then the solution  $Y$  of the transport equation (B.11) satisfies the space-time derivative bounds

$$\sup_{t \in [-T, T]} \llbracket \partial_t^m Y(t, \cdot) \rrbracket_{n, R_Y(t)} \leq \frac{2C_{\mathbf{g}}}{C_{\mathbf{f}} R_{\mathbf{f}}} \frac{(n+m)! m!}{n!} Q^m, \quad (\text{B.20})$$

for all  $m \leq M+1$  and  $n+m \leq N$ , where  $R_Y(t)$  and  $T$  are as defined in (B.14), and

$$Q = \max\{Q_{\mathbf{f}}, Q_{\mathbf{g}}, C_{\mathbf{f}} R_{\mathbf{f}}, 16dC_{\mathbf{f}} R_{\mathbf{g}}\}. \quad (\text{B.21})$$

*Proof of Lemma B.9.* The bound (B.20) is already known to hold when  $m = 0$  due to (B.13) and the definition of  $T$  in (B.14). We next proceed inductively, with respect to  $m$ .

Applying  $\partial_t^m \partial^\alpha$  to (B.11), with  $|\alpha| = n$ , we obtain that

$$\partial_t^{m+1} \partial^\alpha Y = \partial_t^m \partial^\alpha \mathbf{g} - \sum_{j=0}^m \sum_{k=0}^n \sum_{|\beta|=k, \beta \leq \alpha} \binom{m}{j} \binom{\alpha}{\beta} \partial_t^{m-j} \partial^{\alpha-\beta} \mathbf{f} \cdot \nabla \partial_t^j \partial^\beta Y.$$

Taking the supremum over  $|\alpha| = n$ , using the definition (B.1) and the identity (B.2), for any  $R > 0$  we deduce

$$\begin{aligned} \llbracket \partial_t^{m+1} Y \rrbracket_{n,R} &\leq \llbracket \partial_t^m \mathbf{g} \rrbracket_{n,R} + \frac{(n+1)^2}{n! R^n} \sum_{j=0}^m \sum_{k=0}^n \binom{m}{j} \binom{n}{k} \frac{(n-k)! R^{n-k}}{(n-k+1)^2} \llbracket \partial_t^{m-j} \mathbf{f} \rrbracket_{n-k,R} \\ &\quad \times \frac{d(k+1)! R^{k+1}}{(k+2)^2} \llbracket \partial_t^j Y \rrbracket_{k+1,R} \\ &\leq \llbracket \partial_t^m \mathbf{g} \rrbracket_{n,R} + \sum_{j=0}^m \sum_{k=0}^n \binom{m}{j} \frac{(n+1)^2 (k+1) dR}{(n-k+1)^2 (k+2)^2} \llbracket \partial_t^{m-j} \mathbf{f} \rrbracket_{n-k,R} \llbracket \partial_t^j Y \rrbracket_{k+1,R}. \end{aligned}$$

By further appealing to (B.19), choosing  $R = R_Y(t) \geq R_g \geq R_f$ , and using that  $Q \geq Q_g, Q_f$ , we obtain

$$\frac{\llbracket \partial_t^{m+1} Y \rrbracket_{n,R_Y}}{(m+1)! Q^{m+1}} \leq \frac{C_g}{(m+1)Q} + \sum_{j=0}^m \sum_{k=0}^n \frac{d(n+1)^2 (k+1)}{(n-k+1)^2 (k+2)^2 (m+1)} \frac{C_f R_Y}{Q} \frac{\llbracket \partial_t^j Y \rrbracket_{k+1,R_Y}}{j! Q^j}.$$

Since  $0 \leq j \leq m$ , we may appeal to (B.20) inductively, and obtain

$$\frac{\llbracket \partial_t^j Y \rrbracket_{k+1,R_Y}}{j! Q^j} \leq \frac{2C_g (k+1+j)!}{C_f R_f (k+1)!}.$$

Therefore, from the above two inequalities and the bound  $R_Y(t) \leq 2R_g$ , it follows that (B.20) holds at level  $m+1$ , if we are able to establish the bound

$$\frac{C_g}{Q} + \sum_{j=0}^m \sum_{k=0}^n \frac{(n+1)^2}{(n-k+1)^2 (k+2)^2 (m+1)} \frac{2dC_f R_g}{Q} \frac{2C_g (k+j+1)!}{C_f R_f k!} \leq \frac{2C_g (n+m+1)!}{C_f R_f n!}.$$

For this purpose, we first note that for  $0 \leq k \leq n$  and  $0 \leq j \leq m$ , we have that  $\frac{(k+j+1)!}{k!} \leq \frac{(n+m+1)!}{n!}$ . Second, we observe that  $\sum_{j=0}^m \sum_{k=0}^n \frac{(n+1)^2}{(n-k+1)^2 (k+2)^2 (m+1)} \leq 4$ . Thus, it is left to verify the condition

$$\frac{C_f R_f}{Q} + \frac{16dC_f R_g}{Q} \leq 2.$$

In turn, this condition holds due to (B.21). This concludes the inductive proof of (B.20).  $\square$

### B.3 Application to ODEs

Suppose that  $\mathbf{f}: \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  be sufficiently smooth, divergence-free, and let  $X$  be the corresponding flow starting from  $x$ , that is,  $X(\cdot, x)$  satisfies

$$\begin{cases} \partial_t X(t, x) = \mathbf{f}(t, X(t, x)) & \text{in } (-\infty, \infty) \times \mathbb{R}^d, \\ X(0, x) = x, \end{cases} \quad (\text{B.22})$$

and denote by  $X^{-1}(t, \cdot)$  the inverse flow, which thus solves

$$\begin{cases} \partial_t X^{-1} + \mathbf{f} \cdot \nabla X^{-1} = 0 & \text{in } (-\infty, \infty) \times \mathbb{R}^d, \\ X^{-1}(0, x) = x. \end{cases} \quad (\text{B.23})$$

In particular,

$$Y(t, x) = X^{-1}(t, x) - x \quad (\text{B.24})$$

solves (B.11) with  $\mathbf{g} = -\mathbf{f}$ .

**Proposition B.10.** *Suppose that  $\mathbf{f}$  is divergence-free, satisfies the bound (B.12), and let  $X$  be the solution of (B.22). Then, for every  $t \in [-T, T]$ , where  $T$  is as defined in (B.14), we have that*

$$\|\nabla X(t, \cdot) - \mathbf{I}_d\|_{L_x^\infty} \leq |t|dC_{\mathbf{f}}R_{\mathbf{f}} \leq \frac{1}{4}. \quad (\text{B.25})$$

Moreover, for every  $0 \leq n \leq N-1$ , we have that

$$\sup_{t \in [-T, T]} \|\nabla X(t, X^{-1}(t, \cdot))\|_{n, R_{\mathbf{f}}(1+4|t|dC_{\mathbf{f}}R_{\mathbf{f}})^2} \leq (d-1)!(20d)^{d-1}. \quad (\text{B.26})$$

*Proof of Proposition B.10.* The bound (B.25) follows by directly differentiating (B.22), which yields

$$\partial_t(\partial_i X_j - \delta_{ij}) = \partial_k \mathbf{f}^j \circ X \left( \partial_i X^k - \delta_{ik} \right) + \partial_i \mathbf{f}^j \circ X, \quad (\partial_i X_j - \delta_{ij})(0, x) = 0,$$

and applying Grönwall's inequality.

In order to prove (B.26), we recall that

$$\nabla X \circ X^{-1} = (\nabla X^{-1})^{-1} \quad (\text{B.27})$$

as  $d \times d$  matrices. Since  $\mathbf{f}$  is divergence free,  $\det(\nabla X^{-1}) = 1$ , and so  $\nabla X \circ X^{-1}$  is nothing by the transpose of the cofactor matrix associated to  $\nabla X^{-1} = \mathbf{I}_d + \nabla Y$ . In turn, every entry of this cofactor matrix is a sum of  $(d-1)!$  many homogenous monomials of degree  $d-1$  in the entries of the matrix  $\mathbf{I}_d + \nabla Y$ . Since when  $\mathbf{g} = -\mathbf{f}$  the Corollary B.8 yields  $\|\nabla Y(\cdot, t)\|_{n, R_{\mathbf{f}}(1+4tdC_{\mathbf{f}}R_{\mathbf{f}})^2} \leq 4d$ , it follows that  $\|\nabla X^{-1}(\cdot, t)\|_{n, R_{\mathbf{f}}(1+4tdC_{\mathbf{f}}R_{\mathbf{f}})^2} \leq 5d$  for all  $0 \leq n \leq N-1$  and  $d \geq 1$ . Applying Lemma B.1 to a sum of  $(d-1)!$  homogenous monomials of degree  $d-1$  in terms of functions which obey this bound, it follows that

$$\|\nabla X(t, X^{-1}(t, \cdot))\|_{n, R_{\mathbf{f}}(1+4tdC_{\mathbf{f}}R_{\mathbf{f}})^2} \leq (d-1)!4^{d-2}(5d)^{d-1} \leq (d-1)!(20d)^{d-1}$$

for all  $0 \leq n \leq N-1$ .  $\square$

In the course of our proof, we shall also require the regularity of the flow  $\nabla X$  itself, not just its behavior under differentiation when we compose with  $X^{-1}$ .

**Proposition B.11.** *Suppose that  $\mathbf{f}$  satisfies the bound (B.12), and let  $X$  be the solution of (B.22). Then, for every  $t \in [-T, T]$ , where  $T$  is as defined in (B.14), and every  $1 \leq n \leq N-1$ , we have*

$$\sup_{t \in [-T, T]} \|\nabla X(t, x)\|_{n, 8dR_{\mathbf{f}}(1+8dC_{\mathbf{f}}R_{\mathbf{f}}|t|)} \leq 6d. \quad (\text{B.28})$$

*Proof of Proposition B.10.* It turns out that (B.26) is not convenient for setting up an induction scheme. Instead, we will inductively propagate

$$\sup_{t \in [-T, T]} \|X(t, x)\|_{n, 8dR_f(1+B|t|)} \leq (-1)^{n-1} \binom{1/2}{n} \frac{1}{dR_f}, \quad \text{with } B = 8dC_f R_f, \quad (\text{B.29})$$

uniformly for  $t \in [-T, T]$ , for all  $1 \leq n \leq N$ . Without loss of generality we prove (B.29) only for  $0 \leq t \leq T$ . The bound claimed in (B.26) follows from (B.29) since  $4(-1)^{n-1} \binom{1/2}{n} \leq 1/n$  holds for all  $n \geq 1$ , and by the definition of  $T$  we have  $1 + BT \leq 3$ . Indeed,

$$\|X(t, x)\|_{n, 8dR_f(1+B|t|)} \leq \frac{d(n+1)^3}{(n+2)^2} \cdot 24dR_f \cdot \frac{1}{4ndR_f} \leq 6d$$

For  $n = 1$ , we verify (B.29) by using (B.25). This amounts to checking

$$\frac{4(1+|t|dC_f R_f)}{8dR_f(1+B|t|)} \leq \frac{1}{2dR_f} \quad (\text{B.30})$$

which holds since  $B \geq dC_f R_f$ .

Next, we fix  $n \geq 2$  and consider a multi-index  $\beta$  with  $|\beta| = n$ . By Proposition B.2, we have

$$\begin{aligned} \partial_t \partial^\beta X_j &= \partial^\beta X_k (\partial_k \mathbf{f}_j) \circ X + \beta! \sum_{2 \leq |\alpha| \leq n} (\partial^\alpha \mathbf{f}_j) \circ X \sum_{s=1}^n \sum_{p_s(\beta, \alpha)} \prod_{j=1}^s \frac{(\partial^{\ell_j} X)^{\mathbf{k}_j}}{(\mathbf{k}_j!) (\ell_j!)^{\mathbf{k}_j}} \\ &=: \partial^\beta X_k (\partial_k \mathbf{f}_j) \circ X + E_{\text{old}}. \end{aligned} \quad (\text{B.31})$$

In (B.31) we have singled out the term  $\partial^\alpha \mathbf{f}_j$  when  $|\alpha| = 1$ ; this is because only this term comes paired with a derivative of  $X$  which is exactly of order  $n$ ; and indeed one directly verify that for each summand in  $E_{\text{old}}$ , we have  $1 \leq |\ell_j| \leq n-1$ . This expression is thus suitable for an inductive estimate. From (B.12), (B.29), the definition of the partition set in (B.4), the identity provided by Lemma B.3, and the bound given in (B.8), we deduce that

$$\begin{aligned} |E_{\text{old}}| &\leq \beta! \sum_{2 \leq |\alpha| \leq n} \frac{C_f R_f^{|\alpha|} |\alpha|!}{(|\alpha|+1)^2} \sum_{s=1}^n \sum_{p_s(\beta, \alpha)} \prod_{j=1}^s \frac{\left((-1)^{|\ell_j|-1} \binom{1/2}{|\ell_j|} \frac{1}{(|\ell_j|+1)^2} \frac{1}{dR_f} ((8dR_f)(1+Bt))^{|\ell_j|}\right)^{|\mathbf{k}_j|}}{(\mathbf{k}_j!) (\ell_j!)^{|\mathbf{k}_j|}} \\ &\leq C_f \frac{(-1)^n (8dR_f)^n (1+Bt)^n}{(n+1)^2} \beta! \sum_{1 \leq |\alpha| \leq n} (-d)^{-|\alpha|} |\alpha|! \sum_{s=1}^n \sum_{p_s(\beta, \alpha)} \prod_{j=1}^s \frac{\left(\binom{1/2}{|\ell_j|} |\ell_j|!\right)^{|\mathbf{k}_j|}}{(\mathbf{k}_j!) (\ell_j!)^{|\mathbf{k}_j|}} \\ &\leq 2C_f \frac{(-1)^n (8dR_f)^n (1+Bt)^n}{(n+1)^2} (n+1)! \binom{1/2}{n+1}. \end{aligned} \quad (\text{B.32})$$

With (B.32) in hand, we return to (B.31), to which we apply Grönwall's inequality (recall that  $\partial^\beta X|_{t=0} = 0$  since  $|\beta| \geq 2$ ) and deduce that

$$\begin{aligned} \|\partial^\beta X(t, \cdot)\|_{L_x^\infty} &\leq \exp\left(T \|\nabla \mathbf{f}\|_{L_{t,x}^\infty}\right) \int_0^t |E_{\text{old}}(s)| ds \\ &\leq e^{\frac{1}{16}} 2C_f \frac{(-1)^n (8dR_f)^n}{(n+1)^2} (n+1)! \binom{1/2}{n+1} \int_0^t (1+Bs)^n ds \\ &= e^{\frac{1}{16}} 2C_f \frac{(-1)^n (8dR_f)^n}{(n+1)^2} (n+1)! \binom{1/2}{n+1} \frac{(1+Bt)^{n+1}}{B(n+1)}. \end{aligned}$$

The bound (B.29) at level  $n$  now follows once we establish

$$e^{\frac{1}{16}} 2C_f (-1)^n \binom{1/2}{n+1} \frac{1+BT}{B} \leq (-1)^{n-1} \binom{1/2}{n} \frac{1}{dR_f}.$$

Observing that  $(-1) \binom{1/2}{n+1} \binom{1/2}{n}^{-1} = \frac{n-1/2}{n+1} \leq 1$ , and appealing to the definition of  $T$  in (B.14), which gives  $1+BT \leq 1+B/(4dC_f R_f)$ , it follows that the above inequality is implied by the bound

$$\frac{e^{\frac{1}{16}}}{2} \left( 1 + \frac{B}{4dC_f R_f} \right) \leq \frac{B}{4dC_f R_f}.$$

The above estimate now is clearly true by the definition of  $B$  in (B.29). This concludes the proof of the inductive step for (B.29), and thus of the proposition.  $\square$

## C. Ergodic lemmas for periodic functions

### C.1 Some basic ergodic lemmas in one dimension

**Lemma C.1** (Basic  $L^1$  ergodic lemma). *Assume the following:*

- $f : \mathbb{R}^d \rightarrow \mathbb{R}$  is a  $\mathbb{Z}^d$ -periodic function satisfying, for given constants  $C_f > 0$  and  $r \in (0, 1]$ , the quantitative analyticity condition

$$\langle |\nabla^n f| \rangle \leq \frac{C_f n!}{r^n}, \quad \forall n \in \mathbb{N}. \quad (\text{C.1})$$

- $g \in L^1_{\text{loc}}(\mathbb{R}^d)$  is a  $N^{-1}\mathbb{Z}^d$ -periodic function for some given integer  $N \in \mathbb{N}$  with  $N^{-1} \leq r$ .

Then there exists a constant  $C(d) < \infty$  such that

$$|\langle fg \rangle - \langle f \rangle \langle g \rangle| \leq CC_f \langle |g| \rangle \exp\left(-\frac{Nr}{C}\right). \quad (\text{C.2})$$

*Proof of Lemma C.1.* We have that

$$\langle fg \rangle - \langle f \rangle \langle g \rangle = \langle f(g - \langle g \rangle) \rangle = \langle (f - \langle f \rangle)(g - \langle g \rangle) \rangle,$$

where, recall that  $\langle \cdot \rangle$  represents the average on  $[0, 1]^d$ . Using Parseval's identity, and denoting by  $\hat{f}_k$  and  $\hat{g}_k$  the  $k^{\text{th}}$  Fourier series coefficients of  $f$  and  $g$ , respectively, we have

$$\langle (f - \langle f \rangle)(g - \langle g \rangle) \rangle = C \sum_{k \in \mathbb{Z}_*^d} \hat{f}_k \overline{\hat{g}_k}$$

where  $C$  is a dimensional constant (taking into account factors of  $(2\pi)^d$ ), and  $\mathbb{Z}_*^d = \mathbb{Z}^d \setminus \{0\}$ . Moreover, since  $g - \langle g \rangle$  is a zero mean  $N^{-1}\mathbb{Z}^d$ -periodic function, all its *nontrivial* Fourier series coefficients  $\hat{g}_k$  have the property that  $k$  is a nonzero integer multiple of  $N$ . Using this information, and letting  $\tau > 0$ , we thus obtain that

$$|\langle fg \rangle - \langle f \rangle \langle g \rangle| \leq C \sum_{k=N\ell, \ell \in \mathbb{Z}_*^d} e^{\tau|k|} |\hat{f}_k| e^{-\tau|k|} |\hat{g}_k|. \quad (\text{C.3})$$

A simple exercise shows that there exists a constant  $C > 0$  (which only depends on the dimension  $d$ ) such that for  $\tau = r/C$ , the condition (C.1) implies

$$e^{(r/C)|k|} |\widehat{f}_k| \leq CC_f$$

for all  $k \in \mathbb{Z}^d$ . Moreover, we trivially have  $|\widehat{g}_k| \leq C \langle |g| \rangle$  uniformly in  $k$ . Thus, by (C.3)

$$|\langle fg \rangle - \langle f \rangle \langle g \rangle| \leq CC_f \sum_{k=N\ell, \ell \in \mathbb{Z}_*^d} e^{-(r/C)|k|} \leq CC_f \sum_{|\ell|=1}^{\infty} |\ell|^{d-1} e^{-(rN/C)|\ell|}$$

from which (C.2) follows since  $Nr \geq 1$ .  $\square$

**Remark C.2** ( $L^2$  ergodic estimate). We also use the following variant of Lemma C.1: there exists a constant  $C(d) < \infty$  such that if  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  is  $\mathbb{Z}^d$ -periodic and satisfies

$$\langle |\nabla^n f|^2 \rangle^{1/2} \leq \frac{C_f n!}{r^n}, \quad (\text{C.4})$$

and  $g : \mathbb{R}^d \rightarrow \mathbb{R}$  is  $N^{-1}\mathbb{Z}^d$  periodic with  $\langle |g|^2 \rangle < \infty$  and  $Nr \geq 1$ , then we have the bound

$$|\langle |f|^2 |g|^2 \rangle - \langle |f|^2 \rangle \langle |g|^2 \rangle| \leq CC_f^2 \langle |g|^2 \rangle \exp\left(-\frac{Nr}{C}\right). \quad (\text{C.5})$$

Indeed, we may just apply Lemma C.1 with  $f \mapsto f^2$ ,  $g \mapsto g^2$  and  $r \mapsto r/2$ , because the Leibniz rule and assumption (C.1) on  $f$  give

$$\langle |\nabla^n(f^2)| \rangle \leq \sum_{j=0}^n \binom{n}{j} \langle |\nabla^j f|^2 \rangle^{1/2} \langle |\nabla^{n-j} f|^2 \rangle^{1/2} \leq C_f^2 \frac{(n+1)!}{r^n} \leq C_f^2 \frac{n!}{(r/2)^n}.$$

**Lemma C.3** ( $L^1$  and  $L^2$  ergodic lemma with flows). *Assume the following:*

- $X : \mathbb{R}^d \rightarrow \mathbb{R}^d$  is a  $\mathbb{Z}^d$ -periodic, volume-preserving diffeomorphism satisfying, for given constants  $C_X > 0$  and  $R > 0$ , the quantitative analyticity condition

$$\|\nabla^n X\|_{L^\infty(\mathbb{R}^d)} \leq C_X n! R^n, \quad \forall n \in \mathbb{N}. \quad (\text{C.6})$$

- $f : \mathbb{R}^d \rightarrow \mathbb{R}$  is a  $\mathbb{Z}^d$ -periodic function satisfying, for given constants  $C_f > 0$  and  $r \in (0, 1]$ , the quantitative analyticity condition

$$\langle |\nabla^n f| \rangle \leq \frac{C_f n!}{r^n}, \quad \forall n \in \mathbb{N}. \quad (\text{C.7})$$

- $g \in L^1_{\text{loc}}(\mathbb{R}^d)$  is a  $N^{-1}\mathbb{Z}^d$ -periodic function for some given integer  $N \in \mathbb{N}$  satisfying

$$Nr \geq R(r + dC_X). \quad (\text{C.8})$$

Then, there exists  $C(d) < \infty$  such that

$$|\langle f(g \circ X^{-1}) \rangle - \langle f \rangle \langle g \rangle| \leq CC_f \langle |g| \rangle \exp\left(-\frac{Nr}{CR(r + dC_X)}\right). \quad (\text{C.9})$$

Moreover, if we replace (C.7) by the stronger assumption

$$\langle |\nabla^n f|^2 \rangle^{1/2} \leq \frac{C_f n!}{r^n}, \quad (\text{C.10})$$

and the assumption of  $g \in L^1_{\text{loc}}(\mathbb{R}^d)$  by  $g \in L^2_{\text{loc}}(\mathbb{R}^d)$ , then we have the estimate

$$|\langle |f|^2 |g \circ X^{-1}|^2 \rangle - \langle |f|^2 \rangle \langle |g|^2 \rangle| \leq CC_f^2 \langle |g|^2 \rangle \exp\left(-\frac{Nr}{CR(r + dC_X)}\right). \quad (\text{C.11})$$

*Proof of Lemma C.3.* Let  $\tilde{f} = f \circ X$ , which is thus also  $\mathbb{Z}^d$ -periodic and real-analytic. By the assumption (C.1) for  $f$ , (C.6) for  $X$ , and the composition estimate in Proposition B.6, we have

$$\|\nabla^n \tilde{f}\|_{L^\infty(\mathbb{R}^d)} \leq \frac{C_f n!}{\tilde{r}^n}, \quad \text{where} \quad \tilde{r} = \frac{r}{R(r + dC_X)}. \quad (\text{C.12})$$

Now since  $X$  and  $X^{-1}$  are volume preserving, we have that

$$\langle fg \circ X^{-1} \rangle = \langle f \circ X g \rangle = \langle \tilde{f} g \rangle.$$

Thus, by (C.12) we may apply Lemma C.1 to the functions  $\tilde{f}$  and  $g$ , and deduce that

$$|\langle \tilde{f} g \rangle - \langle \tilde{f} \rangle \langle g \rangle| \leq CC_f \langle |g| \rangle \exp\left(-\frac{N\tilde{r}}{C}\right), \quad (\text{C.13})$$

for  $C = C(d)$ , as soon as  $N\tilde{r} \geq 1$ . The bound (C.9) now follows by appealing again to the volume preserving nature of  $X$ , which gives  $\langle \tilde{f} \rangle = \langle f \circ X \rangle = \langle f \rangle$ , and spelling out the  $\tilde{r}$  in (C.12), and using that by assumption we have  $N\tilde{r} \geq 1$ . The proof of (C.11) is the same as the one outlined in Remark C.2, applied to  $g$  and  $\tilde{f}$ .  $\square$

**Lemma C.4** ( $H^{-1}$  ergodic lemma). *Assume that  $f$  and  $g$  satisfying the hypotheses of Remark C.2, and additionally that  $\langle fg \rangle = 0$  and  $\langle g \rangle = 0$ . Then, exists a constant  $C = C(d) < \infty$  such that*

$$\|fg\|_{\dot{H}^{-1}(\mathbb{T}^d)} \leq \frac{C}{N} \langle |f|^2 \rangle^{1/2} \langle |g|^2 \rangle^{1/2} + C_f \langle |g|^2 \rangle^{1/2} \exp\left(-\frac{Nr}{C}\right). \quad (\text{C.14})$$

*Proof of Lemma C.4.* Internally to this proof, for a  $\mathbb{T}^d$  periodic function  $\varphi$  we write  $\mathbb{P}_{\leq L} \varphi$  to denote the truncation of the Fourier series of  $\varphi$  at frequencies  $k \in \mathbb{Z}^d$  such that  $|k| \leq L$ ; accordingly, define  $\mathbb{P}_{>L} \varphi = \varphi - \mathbb{P}_{\leq L} \varphi$ . We use similar notations for  $\mathbb{P}_{<L}$  and  $\mathbb{P}_{\geq L}$ . We will use two useful identities. First, since  $g$  has zero mean and is  $N^{-1}\mathbb{Z}^d$  periodic, then all its active frequencies are nonzero  $\mathbb{Z}^d$ -multiples of  $N$ , and thus

$$g = \mathbb{P}_{\geq N} g. \quad (\text{C.15})$$

Second, using the triangle inequality in the frequency domain combined with (C.15) we may write  $\mathbb{P}_{\leq N/2} f g = \mathbb{P}_{\leq N/2} f \mathbb{P}_{\geq N} g = \mathbb{P}_{\geq N/2} (\mathbb{P}_{\leq N/2} f \mathbb{P}_{\geq N} g)$ . Thus, combining (C.15) and the previous identity, we may rewrite

$$fg = \mathbb{P}_{\geq N/2} (\mathbb{P}_{\leq N/2} f \mathbb{P}_{\geq N} g) + \mathbb{P}_{>N/2} f g. \quad (\text{C.16})$$

In particular,  $\langle fg \rangle = 0$  implies that  $\langle \mathbb{P}_{>N/2} f g \rangle = 0$ .

For brevity of notation, let  $\Lambda = (-\Delta)^{1/2}$ , so that by (C.16) we have

$$\begin{aligned} \|fg\|_{\dot{H}^{-1}(\mathbb{T}^d)} &= \|\Lambda^{-1}(fg)\|_{L^2(\mathbb{T}^d)} \\ &\leq \|\Lambda^{-1} \mathbb{P}_{\geq N/2} (\mathbb{P}_{\leq N/2} f \mathbb{P}_{\geq N} g)\|_{L^2(\mathbb{T}^d)} + \|\Lambda^{-1} (\mathbb{P}_{>N/2} f g)\|_{L^2(\mathbb{T}^d)} \\ &\leq \frac{C}{N} \|\mathbb{P}_{\leq N/2} f g\|_{L^2(\mathbb{T}^d)} + C \|\mathbb{P}_{>N/2} f g\|_{L^2(\mathbb{T}^d)}, \end{aligned} \quad (\text{C.17})$$

where  $C$  is a universal constant (related to the  $2\pi$  which we are not writing anywhere). In the last inequality above we have used two bounds:  $\|\Lambda^{-1}\mathbb{P}_{\geq L}\|_{L^2 \rightarrow L^2} \leq CL^{-1}$ , which is a consequence of Plancherel; and  $\|\Lambda^{-1}\mathbb{P}_{>0}\|_{L^2 \rightarrow L^2} \leq C$ , which holds since for  $\mathbb{Z}^d$  periodic functions we have  $\mathbb{P}_{>0} = \mathbb{P}_{\geq 1}$ .

Let us first inspect the first term on the right side of (C.17). Since  $\nabla^n$  commutes with  $\mathbb{P}_{\leq L}$  and since  $\|\mathbb{P}_{\leq L}\|_{L^2 \rightarrow L^2} \leq 1$ , assumption (C.4) holds with  $f$  replaced by  $\mathbb{P}_{\leq N/2}f$ , with the same constants  $C_f$  and  $r$  (note, we can also replace  $f$  by  $\mathbb{P}_{>N/2}f$ , and this fact will be used later). Thus, we may apply the conclusion of Remark C.2 to the product of  $\mathbb{P}_{\leq N/2}f$  and  $g$ , resulting in

$$\|\mathbb{P}_{\leq N/2}f g\|_{L^2(\mathbb{T}^d)} \leq \left( \|f\|_{L^2(\mathbb{T}^d)} + CC_f \exp(-Nr/C) \right) \|g\|_{L^2(\mathbb{T}^d)}$$

where  $C = C(d)$ . Similarly, we may apply the conclusion of Remark C.2 to the product of  $\mathbb{P}_{>N/2}f$  and  $g$  and deduce

$$\|\mathbb{P}_{>N/2}f g\|_{L^2(\mathbb{T}^d)} \leq \left( \|\mathbb{P}_{>N/2}f\|_{L^2(\mathbb{T}^d)} + CC_f \exp(-Nr/C) \right) \|g\|_{L^2(\mathbb{T}^d)}.$$

Since  $N \geq r^{-1} \geq 1$ , combining the above two displays with (C.17) concludes the proof of (C.14), but only once we show that  $\|\mathbb{P}_{>N/2}f\|_{L^2(\mathbb{T}^d)}$  is exponential small in  $-Nr/C$ .

For this purpose, we note that for any  $\tau > 0$ , we have  $\|\exp(-\tau\Lambda)\mathbb{P}_{>N/2}\|_{L^2 \rightarrow L^2} \leq \exp(-N\tau/C)$ . On the other hand, by expanding the power series of  $\exp$  and appealing to (C.4), we obtain, for every  $0 < \tau < r$ ,

$$\|\exp(\tau\Lambda)f\|_{L^2(\mathbb{T}^d)} \leq \sum_{m \geq 0} \frac{\tau^m}{m!} \|\Lambda^m f\|_{L^2(\mathbb{T}^d)} \leq C_f \sum_{m \geq 0} \frac{\tau^m}{r^m} = \frac{C_f}{1 - \tau/r}$$

Thus, letting  $\tau = r/C$  for a suitable  $C = C(d) \geq 2$ , this paragraph concludes with

$$\|\mathbb{P}_{>N/2}f\|_{L^2(\mathbb{T}^d)} \leq CC_f \exp(-Nr/C),$$

thereby concluding the proof of the Lemma.  $\square$

**Remark C.5** ( $H^{-1}$  ergodic lemma with flows). Let  $f$  and  $g$  be as in Lemma C.4, and let  $X$  be a periodic volume preserving analytic diffeomorphism as in Lemma C.3, with

$$\|\nabla X - \mathbf{I}_d\|_{L^\infty(\mathbb{T}^d)} + \|\nabla X^{-1} - \mathbf{I}_d\|_{L^\infty(\mathbb{T}^d)} \leq 1/2.$$

If  $\langle f g \circ X^{-1} \rangle = \langle g \rangle = 0$ , then in analogy to how (C.5) implies (C.11), from (C.14) we may deduce

$$\|f g \circ X^{-1}\|_{\dot{H}^{-1}(\mathbb{T}^d)} \leq \frac{C}{N} \langle |f|^2 \rangle^{1/2} \langle |g|^2 \rangle^{1/2} + C_f \langle |g|^2 \rangle^{1/2} \exp\left(-\frac{Nr}{CR(r + dC_X)}\right). \quad (\text{C.18})$$

The argument goes as follows. As in the proof of Lemma C.3, define  $\tilde{f} = f \circ X$ , which is thus periodic and satisfies the quantitative analyticity estimates (C.12). By duality, using that  $X$  is volume preserving and  $|\nabla X| \leq 3/2$ , we have

$$\begin{aligned} \|f g \circ X^{-1}\|_{\dot{H}^{-1}(\mathbb{T}^d)} &= \sup_{\varphi \in C^\infty(\mathbb{T}^d), \|\nabla \varphi\|_{L^2} \leq 1} \left| \int f g \circ X^{-1} \varphi \right| \\ &= \sup_{\varphi \in C^\infty(\mathbb{T}^d), \|\nabla \varphi\|_{L^2} \leq 1} \left| \int \tilde{f} g \varphi \circ X \right| \leq \sup_{\tilde{\varphi} \in C^\infty(\mathbb{T}^d), \|\nabla \tilde{\varphi}\|_{L^2} \leq 3/2} \left| \int \tilde{f} g \tilde{\varphi} \right| \leq C \|\tilde{f} g\|_{\dot{H}^{-1}(\mathbb{T}^d)}. \end{aligned}$$

Moreover, note that  $\langle \tilde{f} g \rangle = \langle f g \circ X^{-1} \rangle = 0$  by assumption, so we may directly apply Lemma C.4 to the pair  $\tilde{f}$  and  $g$ , and deduce that (C.18) holds.

**Acknowledgments.** S.A. was supported by NSF grants DMS-1954357 and DMS-2000200 and by the Simons Programme at IHES during a sabbatical visit. V.V. was supported by the NSF CAREER grant DMS-1911413. We thank an anonymous referee for pointing us to the references [Tay53, Ari56] and the connection between the formula (1.13) and Taylor dispersion.

## References

- [ABC14] G. Alberti, S. Bianchini, and G. Crippa. A uniqueness result for the continuity equation in two dimensions. *J. Eur. Math. Soc. (JEMS)*, 16(2):201–234, 2014.
- [ACM19a] G. Alberti, G. Crippa, and A. L. Mazzucato. Exponential self-similar mixing by incompressible flows. *J. Amer. Math. Soc.*, 32(2):445–490, 2019.
- [ACM19b] G. Alberti, G. Crippa, and A. L. Mazzucato. Loss of regularity for the continuity equation with non-Lipschitz velocity field. *Ann. PDE*, 5(1):Paper No. 9, 19, 2019.
- [Aiz78] M. Aizenman. On vector fields as generators of flows: a counterexample to Nelson’s conjecture. *Ann. of Math. (2)*, 107(2):287–296, 1978.
- [AM90] M. Avellaneda and A. J. Majda. Mathematical models with exact renormalization for turbulent transport. *Comm. Math. Phys.*, 131(2):381–429, 1990.
- [AM91] M. Avellaneda and A. J. Majda. An integral representation and bounds on the effective diffusivity in passive advection by laminar and turbulent flows. *Comm. Math. Phys.*, 138(2):339–391, 1991.
- [AM92] M. Avellaneda and A.J. Majda. Mathematical models with exact renormalization for turbulent transport, ii: Fractal interfaces, non-gaussian statistics and the sweeping effect. *Comm. Math. Phys.*, 146(1):139–204, 1992.
- [Amb04] L. Ambrosio. Transport equation and Cauchy problem for  $BV$  vector fields. *Invent. Math.*, 158(2):227–260, 2004.
- [Ari56] R. Aris. On the dispersion of a solute in a fluid flowing through a tube. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 235(1200):67–77, 1956.
- [ARV] S. Armstrong, K. Rowan, and V. Vicol. Anomalous diffusion and uniform Hölder regularity of passive scalars. in preparation.
- [BAO02] G. Ben Arous and H. Owhadi. Super-diffusivity in a shear flow model from perpetual homogenization. *Comm. Math. Phys.*, 227(2):281–302, 2002.
- [BCC<sup>+</sup>22] E. Brue, M. Colombo, G. Crippa, C. De Lellis, and M. Sorella. Onsager critical solutions of the forced Navier-Stokes equations. *arXiv preprint arXiv:2212.08413*, 2022.
- [BCDL21] E. Brue, M. Colombo, and C. De Lellis. Positive solutions of transport equations and classical nonuniqueness of characteristic curves. *Arch. Ration. Mech. Anal.*, 240(2):1055–1090, 2021.
- [BCZ17] J. Bedrossian and M. Coti Zelati. Enhanced dissipation, hypoellipticity, and anomalous small noise inviscid limits in shear flows. *Arch. Ration. Mech. Anal.*, 224(3):1161–1204, 2017.
- [BCZG22] A. Blumenthal, M. Coti Zelati, and R.S. Gvalani. Exponential mixing for random dynamical systems and an example of Pierrehumbert. *arXiv preprint arXiv:2204.13651*, 2022.
- [BDL22] E. Brue and C. De Lellis. Anomalous dissipation for the forced 3D Navier-Stokes equations. *arXiv preprint arXiv:2207.06301*, 2022.
- [BGK98] D. Bernard, K. Gawedzki, and A. Kupiainen. Slow modes in passive advection. *Journal of Statistical Physics*, 90(3):519–569, 1998.
- [BMNV23] T. Buckmaster, N. Masmoudi, M. Novack, and V. Vicol. *Intermittent Convex Integration for the 3D Euler Equations*. Number 217 in Annals of Mathematics Studies. Princeton University Press, 2023.

[BN21] E. Brue and Q.-H. Nguyen. Advection diffusion equations with sobolev velocity field. *Communications in Mathematical Physics*, 383(1):465–487, 2021.

[BV20] T. Buckmaster and V. Vicol. Convex integration and phenomenologies in turbulence. *EMS Surveys in Mathematical Sciences*, 6(1):173–263, 2020.

[CCS22] M. Colombo, G. Crippa, and M. Sorella. Anomalous dissipation and lack of selection in the Obukhov-Corrsin theory of scalar turbulence. *arXiv preprint arXiv:2207.06833*, 2022.

[CDL08] G. Crippa and C. De Lellis. Estimates and regularity results for the DiPerna-Lions flow. *J. Reine Angew. Math.*, 616:15–46, 2008.

[CEIM22] G. Crippa, T. Elgindi, G. Iyer, and A.L. Mazzucato. Growth of Sobolev norms and loss of regularity in transport equations. *Philosophical Transactions of the Royal Society A*, 380(2225):20210024, 2022.

[CET94] P. Constantin, W. E, and E. S. Titi. Onsager’s conjecture on the energy conservation for solutions of Euler’s equation. *Comm. Math. Phys.*, 165(1):207–209, 1994.

[CKRZ08] P. Constantin, A. Kiselev, L. Ryzhik, and A. Zlatoš. Diffusion and mixing in fluid flow. *Ann. of Math. (2)*, 168(2):643–674, 2008.

[CL21] A. Cheskidov and X. Luo. Nonuniqueness of weak solutions for the transport equation at critical space regularity. *Annals of PDE*, 7(1):1–45, 2021.

[Cor51] S. Corrsin. On the spectrum of isotropic temperature fluctuations in an isotropic turbulence. *J. Appl. Phys.*, 22:469–473, 1951.

[CS96] G. M. Constantine and T. H. Savits. A multivariate Faà di Bruno formula with applications. *Trans. Amer. Math. Soc.*, 348(2):503–520, 1996.

[CVW15] P. Constantin, V. Vicol, and J. Wu. Analyticity of Lagrangian trajectories for well posed inviscid incompressible fluid models. *Adv. Math.*, 285:352–393, 2015.

[CZD21] M. Coti Zelati and T.D. Drivas. A stochastic approach to enhanced diffusion. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 22(2):811–834, 2021.

[CZDE20] M. Coti Zelati, M.G. Delgadino, and T.M. Elgindi. On the relation between enhanced dissipation timescales and mixing rates. *Comm. Pure Appl. Math.*, 73(6):1205–1244, 2020.

[DE17] T.D. Drivas and G.L. Eyink. A lagrangian fluctuation–dissipation relation for scalar turbulence. part i. flows with no bounding walls. *Journal of Fluid Mechanics*, 829:153–189, 2017.

[DEIJ22] T.D. Drivas, T.M. Elgindi, G. Iyer, and I.-J. Jeong. Anomalous dissipation in passive scalar transport. *Arch. Ration. Mech. Anal.*, 243(3):1151–1180, 2022.

[Dep03] N. Depauw. Non unicité des solutions bornées pour un champ de vecteurs BV en dehors d’un hyperplan. *C. R. Math. Acad. Sci. Paris*, 337(4):249–252, 2003.

[DL89] R.J. DiPerna and P.-L. Lions. Ordinary differential equations, transport theory and Sobolev spaces. *Invent. Math.*, 98(3):511–547, 1989.

[DSY05] D.A. Donzis, K.R. Sreenivasan, and P.K. Yeung. Scalar dissipation rate and dissipative anomaly in isotropic turbulence. *J. Fluid Mech.*, 532:199–216, 2005.

[EB13] G. L. Eyink and D. Benveniste. Suppression of particle dispersion by sweeping effects in synthetic turbulence. *Phys. Rev. E*, 87:023011, Feb 2013.

[EF11] G. Eyink and U. Frisch. Robert H. Kraichnan. In Peter A. Davidson, Yukio Kaneda, Keith Moffatt, and Katepalli R. Sreenivasan, editors, *A Voyage Through Turbulence*, pages xvi+434. Cambridge University Press, Cambridge, 2011.

[ELM23] T.M. Elgindi, K. Liss, and J.C. Mattingly. Optimal enhanced dissipation and mixing for a time-periodic, Lipschitz velocity field on  $\mathbb{T}^2$ . *arXiv preprint arXiv:2304.05374*, 2023.

[EVE01] W. E and E. Vanden-Eijnden. Turbulent Prandtl number effect on passive scalar advection. *Physica D: Nonlinear Phenomena*, 152:636–645, 2001.

[EZ19] T.M. Elgindi and A. Zlatoš. Universal mixers in all dimensions. *Adv. Math.*, 356:106807, 33, 2019.

[FGV01] G. Falkovich, K. Gawedzki, and M. Vergassola. Particles and fields in fluid turbulence. *Reviews of modern Physics*, 73(4):913, 2001.

[FI19] Y. Feng and G. Iyer. Dissipation enhancement by mixing. *Nonlinearity*, 32(5):1810–1851, 2019.

[FP94] A. Fannjiang and G. Papanicolaou. Convection enhanced diffusion for periodic flows. *SIAM J. Appl. Math.*, 54(2):333–408, 1994.

[Fri95] U. Frisch. *Turbulence*. Cambridge University Press, Cambridge, 1995.

[Gaw08] K. Gawedzki. Soluble models of turbulent transport. *Non-equilibrium statistical mechanics and turbulence*, 355:44–107, 2008.

[GV00] K. Gawedzki and M. Vergassola. Phase transition in the passive scalar advection. *Physica D: Nonlinear Phenomena*, 138(1-2):63–90, 2000.

[IKX14] 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.

[Jab16] P.-E. Jabin. Critical non-Sobolev regularity for continuity equations with rough velocity fields. *J. Differential Equations*, 260(5):4739–4757, 2016.

[JS23] C.J.P. Johansson and M. Sorella. Nontrivial absolutely continuous part of anomalous dissipation measures in time. *arXiv preprint arXiv:2303.09486*, 03 2023.

[KO02] T. Komorowski and S. Olla. On the superdiffusive behavior of passive tracer with a Gaussian drift. *J. Statist. Phys.*, 108(3-4):647–668, 2002.

[KP02] S.G. Krantz and H.R. Parks. *A primer of real analytic functions*. Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks]. Birkhäuser Boston, Inc., Boston, MA, second edition, 2002.

[Kra68] R.H. Kraichnan. Small-scale structure of a scalar field convected by turbulence. *Phys. Fluids*, 11(5):945–953, 1968.

[Kub63] R. Kubo. Stochastic Liouville equations. *J. Math. Phys.*, 4:174–183, 1963.

[Kup03] A. Kupiainen. Nondeterministic dynamics and turbulent transport. In D. Iagolnitzer, V. Rivasseau, and J. Zinn-Justin, editors, *International Conference on Theoretical Physics*, pages 713–726. Birkhäuser Basel, 2003.

[LJR02] Y. Le Jan and O. Raimond. Integration of Brownian vector fields. *The Annals of Probability*, 30(2):826–873, 2002.

[MK99] A. J. Majda and P. R. Kramer. Simplified models for turbulent diffusion: theory, numerical modelling, and physical phenomena. *Phys. Rep.*, 314(4-5):237–574, 1999.

[MS18] S. Modena and L. Székelyhidi. Non-uniqueness for the transport equation with Sobolev vector fields. *Annals of PDE*, 4(2):1–38, 2018.

[Obu49] A.M. Obukhov. The structure of the temperature field in a turbulent flow. *Izvestiya Akad. Nauk SSSR. Ser. Geograf. Geofiz.*, 13:58–69, 1949.

[Pie94] R.T. Pierrehumbert. Tracer microstructure in the large-eddy dominated regime. *Chaos, Solitons & Fractals*, 4(6):1091–1110, 1994.

[PS21] J. Pitcho and M. Sorella. Almost everywhere non-uniqueness of integral curves for divergence-free Sobolev vector fields. *arXiv preprint arXiv:2108.03194*, 2021.

- [Sei13] C. Seis. Maximal mixing by incompressible fluid flows. *Nonlinearity*, 26(12):3279–3289, 2013.
- [SS00] B.I. Shraiman and E.D. Siggia. Scalar turbulence. *Nature*, 405(6787):639–646, 2000.
- [SS10] K.R. Sreenivasan and J. Schumacher. Lagrangian views on turbulent mixing of passive scalars. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 368(1916):1561–1577, 2010.
- [Tay53] G. I. Taylor. Dispersion of soluble matter in solvent flowing slowly through a tube. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 219(1137):186–203, 1953.
- [TD05] D.J. Thomson and B.J. Devenish. Particle pair separation in kinematic simulations. *J. Fluid Mech.*, 526:277–302, 2005.
- [VEE00] E. Vanden-Eijnden and W. E. Generalized flows, intrinsic stochasticity, and turbulent transport. *Proc. Natl. Acad. Sci. USA*, 97(15):8200–8205, 2000.
- [War00] Z. Warhaft. Passive scalars in turbulent flows. In *Annual review of fluid mechanics, Vol. 32*, volume 32 of *Annu. Rev. Fluid Mech.*, pages 203–240. Annual Reviews, Palo Alto, CA, 2000.
- [YZ17] Y. Yao and A. Zlatoš. Mixing and un-mixing by incompressible flows. *J. Eur. Math. Soc. (JEMS)*, 19(7):1911–1948, 2017.