

ANOMALOUS DIFFUSION VIA ITERATIVE QUANTITATIVE HOMOGENIZATION: AN OVERVIEW OF THE MAIN IDEAS

SCOTT ARMSTRONG AND VLAD VICOL

ABSTRACT. Anomalous diffusion is the fundamental ansatz of phenomenological theories of passive scalar turbulence, and has been confirmed numerically and experimentally to an extraordinary extent. The purpose of this survey is to discuss our recent result, in which we construct a class of incompressible vector fields that have many of the properties observed in a fully turbulent velocity field, and for which the associated scalar advection-diffusion equation generically displays anomalous diffusion. Our main contribution is to propose an analytical framework in which to study anomalous diffusion via a backward cascade of renormalized eddy viscosities.

September 3, 2025

1. INTRODUCTION

Let $u = u(t, x) : [0, \infty) \times \mathbb{T}^d \rightarrow \mathbb{R}^d$ be a given divergence-free velocity field, which is \mathbb{T}^d -periodic in space. Here $d \geq 2$ is the space dimension. We are interested in the setting where u is continuous (in space and time), but has *low regularity*; for example, $u \in C_{x,t}^\alpha := C_t^0 C_x^\alpha \cap C_t^\alpha C_x^0$ for some $\alpha \in (0, 1)$. The motivation to consider such vector fields stems from phenomenological theories of fully-developed hydrodynamic turbulence [13, 26, 31, 42, 54, 61, 66, 70]. The Hölder regularity exponent $\alpha = \frac{1}{3}$ is singled out by the Onsager theory of *ideal turbulence* [37, 39], and this exponent plays a prominent role in this manuscript.

For such a given incompressible vector field $u \in C_{x,t}^\alpha$, we consider the advection-diffusion equation

$$\partial_t \theta^\kappa + u \cdot \nabla \theta^\kappa = \kappa \Delta \theta^\kappa, \quad \theta^\kappa|_{t=0} = \theta_{\text{in}}, \quad (1.1)$$

which models the evolution of a *passive scalar* field $\theta^\kappa = \theta^\kappa(t, x) : [0, \infty) \times \mathbb{T}^d \rightarrow \mathbb{R}$ in a background fluid with velocity u . The parameter $\kappa > 0$ represents molecular diffusivity, or, in non-dimensional form, an inverse Péclet number. Since we are interested in the behavior of the fields θ^κ as $\kappa \rightarrow 0$, it is sufficient to consider $\kappa \in (0, 1]$. The initial condition in (1.1) is given by $\theta_{\text{in}} = \theta_{\text{in}}(x)$, which is taken to have zero-mean on \mathbb{T}^d (that is, $\int_{\mathbb{T}^d} \theta_{\text{in}} dx = 0$), and is assumed to be sufficiently smooth (for instance, $\theta_{\text{in}} \in H^1(\mathbb{T}^d)$). The advection-diffusion equation (1.1) is to be solved for $t \geq 0$; for simplicity, here we will only consider $t \in [0, 1]$. We note that if $u \in C_{x,t}^\alpha$ with $\alpha \in (0, 1)$, then for any $\kappa > 0$ the solution θ^κ of (1.1) is *uniquely defined*,¹ and experiences parabolic smoothing: *in positive time*, we have $\theta^\kappa \in C_t^0 C_x^{2,\alpha} \cap C_t^{1,\alpha/2} C_x^0$. Also, note that the solution θ^κ continues to have zero-mean in positive time, namely $\int_{\mathbb{T}^d} \theta^\kappa(t, x) dx = 0$ for all $t \in (0, 1]$.

1.1. A rigorous definition of anomalous diffusion. The fundamental *energy balance* for the advection-diffusion equation is derived by multiplying (1.1) with θ^κ , and using that $\text{div } u = 0$ to write

$$\partial_t \left(\frac{1}{2} |\theta^\kappa|^2 \right) + \text{div}(u \cdot \nabla \theta^\kappa) - \kappa \Delta \left(\frac{1}{2} |\theta^\kappa|^2 \right) + \kappa |\nabla \theta^\kappa|^2 = 0.$$

Integrating the above relation on $[0, t] \times \mathbb{T}^d$ yields the energy balance²

$$\frac{1}{2} \|\theta^\kappa(t, \cdot)\|_{L^2(\mathbb{T}^d)}^2 + \kappa \int_0^t \|\nabla \theta^\kappa(t', \cdot)\|_{L^2(\mathbb{T}^d)}^2 dt' = \frac{1}{2} \|\theta_{\text{in}}\|_{L^2(\mathbb{T}^d)}^2. \quad (1.2)$$

We emphasize that (1.2) may be justified rigorously (with = sign) under a very weak assumption on the incompressible velocity field u , namely $u \in L_t^\infty L_x^d$.

¹At lower regularity than the one considered in this paper, it is possible to construct non-unique weak solutions of the advection-diffusion equation (1.1). This was first achieved in [23, 57, 58], by treating the diffusion term as an error in an “intermittent convex-integration scheme” for the transport equation; this idea was introduced in [17, 18] for the 3D Navier-Stokes equations. In the works [23, 57, 58], the velocity field u and the scalar field θ^κ are constructed simultaneously, resulting in low regularity for both. See also the recent work [59] for a non-uniqueness proof based on stochastic Lagrangians.

²The quantity $\|\theta^\kappa\|_{L^2(\mathbb{T}^d)}^2$ is the passive scalar’s *variance* (the mean of θ^κ is fixed to equal zero), but we may also think about it as an *energy*.

The uniform-in- κ a priori estimate on θ^κ in $L_t^\infty L_x^2$ provided by (1.2) shows that sequences $\{\theta^{\kappa_j}\}_{j \geq 0}$, with $\kappa_j \rightarrow 0$ as $j \rightarrow \infty$, have sub-sequential weak-* limits in $L_t^\infty L_x^2$. Moreover, the linearity of (1.1) implies that any such weak-* limit point $\bar{\theta}$ is a $L_t^\infty L_x^2$ weak solution of the transport equation

$$\partial_t \bar{\theta} + \operatorname{div}(u \bar{\theta}) = 0, \quad \bar{\theta}|_{t=0} = \theta_{\text{in}}. \quad (1.3)$$

Since under weak convergence we have $\|\bar{\theta}(t, \cdot)\|_{L^2(\mathbb{T}^d)} \leq \liminf_{\kappa \rightarrow 0} \|\theta^\kappa(t, \cdot)\|_{L^2(\mathbb{T}^d)}$ for $t \in [0, 1]$, with (1.2) we obtain

$$\begin{aligned} 0 \leq \limsup_{\kappa \rightarrow 0} \kappa \int_0^t \|\nabla \theta^\kappa(t', \cdot)\|_{L^2(\mathbb{T}^d)}^2 dt' &= \frac{1}{2} \|\theta_{\text{in}}\|_{L^2(\mathbb{T}^d)}^2 - \liminf_{\kappa \rightarrow 0} \frac{1}{2} \|\theta^\kappa(t, \cdot)\|_{L^2(\mathbb{T}^d)}^2 \\ &\leq \frac{1}{2} \|\theta_{\text{in}}\|_{L^2(\mathbb{T}^d)}^2 - \frac{1}{2} \|\bar{\theta}(t, \cdot)\|_{L^2(\mathbb{T}^d)}^2 \end{aligned} \quad (1.4)$$

for $t \in [0, 1]$. The right side of (1.4) depends only on the solution $\bar{\theta}$ of the transport equation (1.3) which arose in the weak-* limit.³ The bound on the energy dissipation rate obtained in (1.4) shows that if all $L_t^\infty L_x^2$ weak solutions of the transport equation (1.3) have constant-in-time $L^2(\mathbb{T}^d)$ -norm, then the right side of (1.4) vanishes identically, and thus $\limsup_{\kappa \rightarrow 0} \kappa \int_0^t \|\nabla \theta^\kappa(t', \cdot)\|_{L^2(\mathbb{T}^d)}^2 dt' = 0$. Such a situation occurs for instance if the solution of (1.3) is uniquely obtained via ODE theory under the Cauchy-Lipschitz assumption $u \in L_t^1 W_x^{1,\infty}$. At a lower space integrability of the velocity gradient, the Di Perna-Lions theory [30] guarantees that if $u \in L_t^1 W_x^{1,1}$, then all $L_t^\infty L_x^2$ weak solutions of (1.3) are *renormalized* and thus they conserve their energy as a function of time.⁴

The above discussion shows that if $\nabla u \in L_{x,t}^1$, then by the Di Perna-Lions theory we automatically have $\limsup_{\kappa \rightarrow 0} \kappa \int_0^t \|\nabla \theta^\kappa(t', \cdot)\|_{L^2(\mathbb{T}^d)}^2 dt' = 0$ for $t \in [0, 1]$. Turbulent flows u however do not possess the property that $\nabla u \in L_{x,t}^1$; instead, we expect u to have Onsager supercritical regularity, such as $u \in C_{x,t}^\alpha$ for some $\alpha \leq \frac{1}{3}$ (see [19, 37, 39, 42] and the references therein). This leaves open the possibility that the right side of (1.4) does not vanish identically, motivating the following definition:

Definition 1.1 (Anomalous diffusion). Fix an incompressible vector field $u \in C_{x,t}^0$. If for any $\theta_{\text{in}} \in H^1(\mathbb{T}^d)$ with $\int_{\mathbb{T}^d} \theta_{\text{in}}(x) dx = 0$, there exists $0 < \varrho \leq \frac{1}{2}$ such that the family of unique solutions $\{\theta^\kappa\}_{\kappa > 0}$ of the Cauchy problem (1.1) satisfy

$$\limsup_{\kappa \rightarrow 0} \kappa \int_0^1 \int_{\mathbb{T}^d} |\nabla \theta^\kappa(t', x)|^2 dx dt' \geq \varrho \|\theta_{\text{in}}\|_{L^2(\mathbb{T}^d)}^2, \quad (1.5)$$

we say the the advection-diffusion equation associated to u exhibits *anomalous diffusion* on $[0, 1]$; equivalently, *anomalous dissipation of the passive scalar's variance*.⁵

The anomalous diffusion associated to the time interval $[0, 1]$, as defined in Definition 1.1, may equivalently be introduced for any time interval $[0, t]$ with $t \in (0, 1]$. Moreover, given any sequence $\kappa_n \rightarrow 0$, one may find a further subsequence $\kappa_{n_j} \rightarrow 0$ along which the *time dissipation measure* $\mathcal{E}(dt)$ is the weak-* limit in the sense of bounded measures

$$\mathcal{E}(dt) := \lim_{j \rightarrow \infty} \kappa_{n_j} \|\nabla \theta^{\kappa_{n_j}}(t, \cdot)\|_{L^2(\mathbb{T}^d)}^2. \quad (1.6)$$

Here we abuse notation, as clearly \mathcal{E} also depends on the initial datum θ_{in} , and also on the chosen sub-sequence $\{\kappa_{n_j}\}_{j \geq 0}$. With (1.4) and (1.6), relation (1.5) becomes $\varrho \|\theta_{\text{in}}\|_{L^2(\mathbb{T}^d)}^2 \leq \int_0^1 \mathcal{E}(dt) \leq \frac{1}{2} \|\theta_{\text{in}}\|_{L^2(\mathbb{T}^d)}^2$.

1.2. Predictions from hydrodynamic turbulence. It is widely accepted that if the vector field u represents a homogenous isotropic turbulent velocity field, then solutions of (1.1) exhibit anomalous diffusion in the sense of Definition 1.1. This prediction was first discussed by Obukhov [61], who drew direct analogies to the anomalous dissipation of kinetic energy in a turbulent fluid, postulated by Kolmogorov's 1941 phenomenological theory [51]. Compelling numerical and experimental evidence for anomalous diffusion is presented in [31, 66, 67, 70]. In this setting, one expects that the time dissipation measure $\mathcal{E}(dt)$ is non-atomic, and is maybe even absolutely continuous with respect to Lebesgue measure on $[0, 1]$.

³For vector fields u whose regularity lies *below* $L_t^1 W_x^{1,\infty}$, e.g. for $u \in L_t^1 C_x^\alpha$ with $\alpha < 1$, the transport equations are expected to have non-unique solutions. Different subsequences $\kappa_j \rightarrow 0$ could in principle select different such solutions (see Theorem 1.2 below). Therefore, (1.4) should be written in terms of sequences $\kappa_j \rightarrow 0$, and for a *given* weak-* limit $\bar{\theta}$.

⁴See also the work of Ambrosio [2] for $u \in L_t^1 \text{BV}_x$.

⁵For passive scalar transport, anomalous diffusion (1.5) is equivalent to Lagrangian spontaneous stochasticity [32]. Roughly speaking, this means that backwards-in-time transition probabilities of the SDE used to represent the advection-diffusion equation, do not converge as $\kappa \rightarrow 0$ to a Dirac mass around a deterministic Lagrangian trajectory.

Under the ansatz that the advection-diffusion equation (1.1) associated to u exhibits anomalous diffusion (a so-called “experimental fact”), and assuming that u represents a homogenous isotropic turbulent velocity field exhibiting *monofractal scaling in the inertial range*, Obukhov [61] and Corrsin [26] have furthermore predicted that the sequence $\{\theta^\kappa\}_{\kappa>0}$ retains uniform-in- κ Hölder regularity, with exponents which may be determined from scaling.

In modern mathematical terms, in analogy with the Onsager conjecture for the incompressible Euler⁶ equations [19, 25, 28, 29, 36, 39],⁷ one may propose a *mathematical idealization* of the Obukhov-Corrsin prediction, and postulate the dichotomy:

- (i) If $u \in C_{x,t}^\alpha$ for some $\alpha \in (0, 1)$ is divergence-free, and if the family of solutions $\{\theta^\kappa\}_{\kappa>0}$ of (1.1) are uniformly-in- κ bounded in $L_t^2 C_x^{\bar{\alpha}}$ for some $\bar{\alpha} > \frac{1-\alpha}{2}$, then there is no anomalous diffusion, namely $\lim_{\kappa \rightarrow 0} \kappa \|\nabla \theta^\kappa\|_{L_t^2 L_x^2}^2 = 0$.
- (ii) For each $\alpha \in (0, \frac{1}{3})$,⁸ there exists divergence-free $u \in C_{x,t}^\alpha$, such that anomalous diffusion holds in the sense of Definition 1.1, with a non-atomic time-dissipation measure $\mathcal{E}(dt)$. Moreover, for *all* smooth initial conditions θ_{in} , the solutions $\{\theta^\kappa\}_{\kappa>0}$ of (1.1) are uniformly-in- κ bounded in $L_t^2 C_x^{\bar{\alpha}}$ for $\bar{\alpha} < \frac{1-\alpha}{2}$.

Part (i) of the Obukhov-Corrsin dichotomy is known, and follows directly from the commutator estimate of Constantin, E, and Titi [25] (see e.g. [33] or [24]).⁹ As stated, part (ii) of the dichotomy remains open (see §1.3 for partial progress). The main purpose of our work [9] is to provide a new perspective on anomalous diffusion for the passive scalar equation, in the hope that these ideas will lead to a resolution of part (ii) of the Obukhov-Corrsin dichotomy.

1.3. Previous results. From a rigorous mathematical perspective, prior to our work [9], we are aware of only two (classes of) examples of incompressible vector fields u for which the associated passive scalar equation (1.1) exhibits (a form of) anomalous diffusion (1.5): the first is the stochastic Kraichnan model (introduced by Kraichnan in [52]), while the second is a class of deterministic singularly-focusing mixing flows (introduced by Drivas, Elgindi, Iyer, and Jeong in [33]). We next discuss these examples, but do so *exceedingly briefly*. For a complete list of references and appropriate context, we refer the interested readers to the introductions of [9, 20, 32, 33, 37, 55].

Kraichnan proposed in [52] a model in which $u = u^\nu$ is a realization of a homogenous isotropic stationary zero-mean Gaussian random field, which is delta-correlated in time, and is colored in space in order to match the Kolmogorov 1941 phenomenological theory: the space covariance is fully-smooth at length scales $\lesssim \nu^{\frac{1}{2}}$ (a UV cutoff serving as a proxy for the Kolmogorov dissipation scale), and behaves like a C^α function at length scales $\gtrsim \nu^{\frac{1}{2}}$ (a synthetic power spectrum in the inertial range). For such a stationary Gaussian random field u^ν , instead of (1.1) one studies the Stratonovich SDE $d\theta^{\kappa,\nu} + u^\nu \odot d\theta^{\kappa,\nu} = \kappa \Delta \theta^{\kappa,\nu} dt$. The main result is that in the joint limit $\nu, \kappa \rightarrow 0$, and for a.e. realization of the field u^ν , and any $\theta_{\text{in}} \in L^2(\mathbb{T}^d)$ of zero-mean, anomalous diffusion holds; see Bernard, Gawedski, and Kupiainen [14], Le Jan and Raimond [53], or Falkovich, Gawedski, and Vergassola [40]. A fully rigorous, complete, and concise proof of anomalous diffusion for the Kraichnan model (and its variants) may be found in the recent work of Rowan [62].

The main drawback of the Kraichnan model stems from the white-noise temporal correlation of the vector field u , which is indeed so rough that it is likely entirely responsible for the anomalous diffusivity; see the mathematical discussion in [62] and the numerical comparisons made in [68].

The first deterministic vector field u for which the dissipation anomaly (1.5) can be established rigorously, was constructed by Drivas, Elgindi, Iyer, and Jeong [33]. These authors show that for any $\alpha \in [0, 1)$ and any dimension $d \geq 2$, there exists $u \in L^1([0, 1]; C^\alpha(\mathbb{T}^d)) \cap L^\infty([0, 1]; L^\infty(\mathbb{T}^d))$, such that $u(t, \cdot)$ is smooth for any $t < 1$, and such that for any $\theta_{\text{in}} \in H^2(\mathbb{T}^d)$ which is sufficiently close (in L^2) to a an eigenfunction of the Laplacian, anomalous diffusion (1.5) holds, with a purely-atomic time dissipation measure $\mathcal{E}(dt) = \rho(\theta_{\text{in}}) \delta_{t=1}$. The construction in [33] is based on “quasi-self-similarly” (as $t \rightarrow 1^-$) accelerating an inviscid mixing dynamics [1], and treating the diffusion term as a perturbation. This is made possible by the singular nature of the “speed-up” of u with respect to time, and the fact that at each time $t < 1$ the field u has only one “active scale” with respect to space.

⁶A weak solution $(u, p) \in L_{x,t}^2 \times L_{x,t}^1$ of the incompressible Euler equations solves the system $\partial_t u + \operatorname{div}(u \otimes u) + \nabla p = 0$, with $\operatorname{div} u = 0$, in the sense of space-time distributions.

⁷For the incompressible Euler equations, the Onsager dichotomy is a “Theorem”, except for the most important endpoint case $\alpha = \frac{1}{3}$; see [22, 34] for the sharpest *rigidity* statements, and [16, 44, 45, 47, 60] for the latest *flexibility* statements in the Onsager Theorem.

⁸The reason that we do not include here the full range $\alpha \in (0, 1)$ is discussed in §4 below.

⁹Specifically, one may modify the mollification+commutator argument of [25] to establish the bound $\kappa \int_0^1 \|\nabla \theta^\kappa(\cdot, t)\|_{L^2}^2 dt \lesssim \kappa^{\frac{\alpha+2\bar{\alpha}-1}{\alpha+1}} (1 + \|u\|_{L_t^\infty C_x^\alpha})^{\frac{2(1-\bar{\alpha})}{\alpha+1}} \|\theta^\kappa\|_{L_t^\infty C_x^{\bar{\alpha}}}^2$.

The ideas in [33] proved to be quite popular, leading to a number of further refinements [15, 21, 24, 35, 49, 50, 63]. Notably, Colombo, Crippa, and Sorella [24] gave examples of a smooth initial conditions θ_{in} , and of a divergence-free vector field $u \in C_t^0 C_x^\alpha$ (for any $\alpha < 1$), such that the associated family of solutions $\{\theta^\kappa\}_{\kappa>0}$ is uniformly-in- κ bounded in $L_t^2 C_x^{\bar{\alpha}}$ for any $\bar{\alpha} < \frac{1-\alpha}{2}$, consistent with the Obukhov-Corrsin dichotomy; moreover, (1.5) holds with $\mathcal{E}(dt) = \rho(\theta_{in})\delta_{t=1}$, and sub-sequential limits select different weak solutions of the transport equation.¹⁰

This result was further refined by Elgindi and Liss [35], who for any $\alpha \in (0, 1)$ provide a divergence-free vector field $u \in C^\infty([0, 1]; C^\alpha(\mathbb{T}^d))$ such that for *all* zero-mean and smooth initial conditions θ_{in} (instead of just *one initial condition*) anomalous diffusion holds with $\mathcal{E}(dt) = \rho(\theta_{in})\delta_{t=1}$; moreover, $\{\theta^\kappa\}_{\kappa>0}$ is uniformly-in- κ bounded in $L_t^2 C_x^{\bar{\alpha}}$ for any $\bar{\alpha} < \frac{1-\alpha}{2}$. We also mention the work of Johannson-Sorella [50] who construct an *autonomous* divergence-free velocity field $u \in C_x^\alpha$ and a smooth zero-mean initial condition θ_{in} , such that (1.5) holds; the associated time dissipation measure $\mathcal{E}(dt)$ is purely absolutely continuous w.r.t. the Lebesgue on $[0, 1]$. Lastly, we note the recent paper of Hess-Childs and Rowan [63] who for any $\alpha \in (0, 1)$ construct $u \in C_t^0 C_x^\alpha \cap C_t^{\frac{\alpha}{1-\alpha}} C_x^0$ such that for any zero-mean $\theta_{in} \in L^1(\mathbb{T}^2)$, the passive scalar θ^κ exhibits *asymptotic total dissipation* as $\kappa \rightarrow 0$, at $\{t = 1\}$; that is, $\lim_{\kappa \rightarrow 0} \|\theta^\kappa(1, \cdot)\|_{L^1(\mathbb{T}^2)} = 0$.

The main drawbacks of the constructions in [24, 33, 35] are that the anomalous dissipation occurs at only a discrete set of times ($\mathcal{E}(dt)$ is proportional to $\delta_{t=1}$), the vector field u has only one active scale at each time $t \in [0, 1]$, and the advection-diffusion equation (1.1) is treated as a perturbation of the transport equation (1.3). Taken together, these properties are inconsistent with the observed properties of turbulent flows.

Note that the result in [35] *nearly solves* part (ii) of the Obukhov-Corrsin dichotomy discussed above, except for the singular nature of the time dissipation measure. In §4 we argue that this seemingly minor caveat (the atomic nature of $\mathcal{E}(dt)$), combined with the fact that anomalous diffusion is meant to hold for all smooth zero-mean initial conditions θ_{in} , is in fact fundamental to the Obukhov-Corrsin dichotomy, and we present a conjecture in this direction (see Conjecture 4.1).

1.4. Main result. Our goal in [9] was to build “minimally realistic examples” of “fluid-like” velocity fields u , for which anomalous diffusion holds in the sense of Definition 1.1,¹¹ and such that the time dissipation measure $\mathcal{E}(dt)$ defined in (1.6) is *non-atomic*.¹²

From a broader perspective, our goal was to highlight what we believe to be the “right reason” for which anomalous diffusion holds: it is not that the underlying transport dynamics exhibits *mixing*, which is accelerated towards a discrete set of times; rather, we show that as one “zooms out” from small scales to large scales, one may observe a backwards cascade of renormalized “eddy diffusivities” taking shape, in such a way that the large scale motion experiences $O(1)$ diffusivity independently of the vanishing molecular diffusivity (as $\kappa \rightarrow 0$). Of course, this picture is not at all original and is indeed widely shared by physicists, where it goes under the name *renormalization group*: see for instance Frisch [42, Section 9.6]. Our main result is:

Theorem 1.2 (Theorem 1.1 in [9]). *Fix $d \geq 2$ and $\alpha \in (0, \frac{1}{3})$.¹³ There exists an incompressible vector field $u \in C^0([0, 1]; C^\alpha(\mathbb{T}^d)) \cap C^\alpha([0, 1]; C^0(\mathbb{T}^d))$, such that for every $\theta_{in} \in \dot{H}^1(\mathbb{T}^d)$, the family of solutions $\{\theta^\kappa\}_{\kappa>0}$ of the advection-diffusion equation (1.1) with velocity u , exhibits anomalous diffusion in the sense of Definition 1.1, with a non-atomic time-dissipation measure. Moreover, we have:*

- The parameter $\varrho \in (0, \frac{1}{2}]$ appearing in (1.5) depends only on α , d , and a lower bound on the natural length scale of the initial datum, given by $L_{in} := \frac{\|\theta_{in}\|_{L^2}}{\|\nabla \theta_{in}\|_{L^2}}$. Specifically, for any $\varepsilon > 0$, we may find $c = c(d, \alpha, \varepsilon) > 0$ such that $\varrho = c L_{in}^{\frac{2(1+\alpha)}{1-\alpha} + \varepsilon}$.
- There exists a sequence $\kappa_j \rightarrow 0$, which depends on α and d but is independent of θ_{in} , along which the lower bound in (1.5) is realized; upon defining $\mathcal{K} = \cup_{j \geq 1} [\frac{1}{2}\kappa_j, 2\kappa_j]$, we have $\inf_{\kappa \in \mathcal{K}} \kappa \|\nabla \theta^\kappa\|_{L^2((0,1) \times \mathbb{T}^d)}^2 \geq \varrho \|\theta_{in}\|_{L^2(\mathbb{T}^d)}^2$.
- There exists $\mu > 0$, such that for any sequence $\kappa_i \rightarrow 0$ with $\{\kappa_i\}_{i \geq 1} \subset \mathcal{K}$, the sequence of solutions $\{\theta^{\kappa_i}\}_{i \geq 1}$ is uniformly-in- κ_i bounded in $C^{0,\mu}([0, 1]; L^2(\mathbb{T}^d))$, as $\kappa_i \rightarrow 0$. In particular, the anomalous dissipation occurs continuously in time: along this subsequence we have that $\int_0^t \mathcal{E}(dt')$ is Hölder continuous in $t \in [0, 1]$. See [9, Remark 5.4].

¹⁰At the lower regularity $u \in L_{x,t}^\infty$, the vanishing viscosity limit fails to select not just unique solutions of the transport equation, but also physically admissible ones, in the sense of non-increasing energy/entropy. [46]

¹¹In particular, (1.5) holds for *all* zero-mean smooth initial conditions.

¹²In particular, the anomalous dissipation occurs continuously in time.

¹³The restriction $\alpha < \frac{1}{3}$ is intimately related to Conjecture 4.1 below.

- There exists sequences $\{\kappa_i\}_{i \geq 1}, \{\kappa'_i\}_{i \geq 1} \subset \mathcal{K}$ such that as $\kappa_i, \kappa'_i \rightarrow 0$, the sequences $\{\theta^{\kappa_i}\}_{i \geq 1}$ and $\{\theta^{\kappa'_i}\}_{i \geq 1}$ converge strongly in $C_t^0 L_x^2$ to two distinct bounded weak solutions $\bar{\theta}$ and $\bar{\theta}'$ of the transport equation (1.3). This demonstrates the lack of a selection principle for vanishing diffusivity limits of the advection-diffusion equation towards the transport equation. See [9, Proposition 5.5].
- The constructed vector field u is “nearly” a weak solution of the incompressible Euler equations, in the sense of [9, Proposition 2.8].
- The class of vector fields u for which the all the conclusions of this Theorem hold is dense in the class of all smooth zero-mean divergence-free vector fields, with respect to the $C_{x,t}^0$ topology. See Remark 3.1 below.

The construction of the vector field u postulated in Theorem 1.2 has many features in common with Nash-type convex-integration schemes for the incompressible Euler equations: multi-scale constructions with super-geometric scale separation, infinitely-many nested active scales with “building blocks” that are stationary Euler solutions, and the usage of Lagrangian flows to advect small-scale features by large-scale flows. On the other hand, the analysis of the passive scalars θ^κ is based on ideas from quantitative homogenization. Specifically, we need to perform homogenization *iteratively across many length scales*. In this way, we observe the emergence of a sequence of effective “eddy” diffusivities, which are recursively defined by homogenizing “up the scales”, from the small scale at which molecular diffusion acts, to the large scale of the periodic box. In this process, every homogenization step resembles Taylor [69] diffusion specialized to the case of a shear flow. An outline of the proof of Theorem 1.2 is provided in §3 below.

1.5. Anomalous diffusion via weak solutions of 3D incompressible Euler. We wish to highlight a very important extension of Theorem 1.2, which was recently obtained by Burczak, Székelyhidi, and Wu in [20]. Whereas the vector field u in Theorem 1.2 may only be shown to be “nearly” a weak solution of the incompressible Euler equations on \mathbb{T}^d ,¹⁴ the authors of [20] have shown that anomalous dissipation (as in Theorem 1.2) holds for large classes of vector fields $u \in C_{x,t}^\alpha$ (with $\alpha < \frac{1}{3}$), that are true weak solutions of three-dimensional incompressible Euler equations:

Theorem 1.3 (Burczak, Székelyhidi, Wu [20]). *Fix $d = 3$. For any $\alpha \in (0, \frac{1}{3})$, there exists a weak solution $u \in C^0([0, 1]; C^\alpha(\mathbb{T}^3)) \cap C^\alpha([0, 1]; C^0(\mathbb{T}^3))$ of the incompressible Euler equations, such that the advection-diffusion equation (1.1) with velocity u exhibits anomalous diffusion, precisely as in Theorem 1.2. Additionally, the family of all such weak solutions $u \in C_{x,t}^\alpha$ of the Euler equations is dense in the class of all solenoidal vector fields, with respect to the $C_t^0 H_x^{-1}$ topology.*

The proof of Theorem 1.3 uses the concepts and analytical methodology of our work [9], and elegantly combines them with ideas from convex-integration / Nash-schemes for the incompressible Euler equations (the technical implementation in [16] is of particular relevance). Notably, Burczak, Székelyhidi, and Wu succeed in [20] to interlace the backwards energy cascade characteristic of convex integration schemes, with the backwards cascade of renormalized diffusivities which is underpinning iterated homogenization. At a technical level, [20] needs to contend with the fact that the correctors are not anymore explicit,¹⁵ and with the fact that Isett’s gluing technique [47] necessitates a more involved time-homogenization step.¹⁶ Lastly, we note that the proof in [20] necessitates at least three space dimensions, due to the usage of non-intersecting Mikado flows [27]. It is conceivable that a two-dimensional result is attainable by combining the ideas in [9, 20] with the Newton-Nash scheme which was used to resolve the 2D Onsager conjecture [45].

2. BACKGROUND ON QUANTITATIVE HOMOGENIZATION

Before turning to the proof of Theorem 1.2, it is instructive to briefly recall the basic two-scale expansion in parabolic periodic homogenization, emphasizing the relation between the effective equation and the phenomenon of *advection-enhanced diffusion* (see §2.1), and then to briefly recall some of the basic themes of *quantitative homogenization*, highlighted in the context of the homogenization problem of a periodic shear flow, and leading to the Taylor diffusion formula (see §2.2).

¹⁴See Proposition 2.8 in [9].

¹⁵As opposed to [9], the solution of the cell-problem in [20] is not explicit. Instead, the correctors are very well-approximated by explicit functions. This is known to be sufficient, because the space-homogenization step is stable under small perturbations in the diffusion matrix.

¹⁶Time-homogenization due to switching of flow-maps was a difficulty already faced and resolved in [9]. The complication in [20] stems from the turning-on and turning-off of the fast oscillations in u , due to the so-called “gluing step”.

2.1. The two-scale expansion in parabolic homogenization. We begin with a general presentation of the parabolic two-scale expansion in the periodic setting; this material is standard (see for example [43, 48, 64]). We consider a $\mathbb{Z} \times \mathbb{Z}^d$ -periodic, mean-zero, incompressible vector field $b = b(t, x)$ and for $\varepsilon, \kappa > 0$, consider the advection-diffusion equation

$$\partial_t \theta_\varepsilon - \kappa \Delta \theta_\varepsilon + \frac{1}{\varepsilon} b\left(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon}\right) \cdot \nabla \theta_\varepsilon = 0. \quad (2.1)$$

By introducing a stream matrix s for the vector field b (that is, an anti-symmetric matrix such that $-\operatorname{div} s = b$), the advection term in (2.1) may be expressed as a second-order term $\frac{1}{\varepsilon} b\left(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon}\right) \cdot \nabla \theta_\varepsilon = -\operatorname{div}(s\left(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon}\right) \nabla \theta_\varepsilon)$; in turn, this allows us to write (2.1) as

$$\partial_t \theta_\varepsilon - \operatorname{div}(a^\varepsilon \nabla \theta_\varepsilon) = 0, \quad \text{where } a^\varepsilon(t, x) := \kappa I_d + s\left(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon}\right) = a\left(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon}\right). \quad (2.2)$$

Note that (2.2) contains a diffusion matrix a^ε that oscillates fast (it depends on ε), and is not symmetric (since s is skew-symmetric). Homogenization theory quantifies the convergence as $\varepsilon \rightarrow 0$ (say with respect to the $L^2_{t,x}$ topology) of solutions θ_ε of (2.2) to solutions θ of the *effective equation*

$$\partial_t \theta - \operatorname{div}(\bar{a} \nabla \theta) = 0, \quad (2.3)$$

where:

- the homogenized/effective diffusion matrix \bar{a} is given by

$$\bar{a}e = \langle\langle a(e + \nabla \chi_e) \rangle\rangle, \quad e \in \mathbb{S}^{d-1}, \quad a := \kappa I_d + s, \quad (2.4)$$

- $\langle\langle \cdot \rangle\rangle$ denotes the space-time average of a $\mathbb{Z} \times \mathbb{Z}^d$ -periodic function,
- χ_e is the *corrector* with slope e ; that is, the unique $\mathbb{Z} \times \mathbb{Z}^d$ -periodic solution of the *cell problem*

$$\partial_t \chi_e - \operatorname{div}(a(e + \nabla \chi_e)) = 0, \quad \langle\langle \chi_e \rangle\rangle = 0. \quad (2.5)$$

The above definitions show that the symmetric part of the effective diffusion matrix \bar{a} is given by

$$\frac{1}{2}(\bar{a} + \bar{a}^t)_{ij} = \kappa \delta_{ij} + \kappa \langle\langle \nabla \chi_{e_i} \cdot \nabla \chi_{e_j} \rangle\rangle. \quad (2.6)$$

The second term on the right side of (2.6) is positive (as a nonnegative definite matrix), and therefore the symmetric part of \bar{a} is larger than the original diffusion matrix κI_d . This effect is called the *enhancement of diffusivity due to advection*.¹⁷

The convergence $\|\theta_\varepsilon - \theta\|_{L^2_{t,x}} \rightarrow 0$ as $\varepsilon \rightarrow 0$, where θ_ε solves the microscopic diffusion equation (2.2) and θ is an appropriately chosen solution of the macroscopic or effective equation (2.3) is summarized informally as:

“the operator $\partial_t - \operatorname{div}(a^\varepsilon \nabla)$ homogenizes as $\varepsilon \rightarrow 0$ to the operator $\partial_t - \operatorname{div}(\bar{a} \nabla)$.”

A standard formalization of the above statement is achieved by proving that θ_ε is well-approximated by its “two-scale expansion”, which is defined in (2.7) below.

The *parabolic* two-scale expansion is constructed as follows. In addition to the corrector with slope e , χ_e (defined in (2.5)), and to the homogenized matrix \bar{a} (defined in (2.4)), we introduce:

- the space-homogenized matrix $\bar{a}(t)$ is defined for $t \in \mathbb{R}$ by

$$\bar{a}(t)e = \langle a(t, \cdot)(e + \nabla \chi_e(t, \cdot)) \rangle, \quad e \in \mathbb{S}^{d-1},$$

where $\langle \cdot \rangle$ denotes the space average of a \mathbb{Z}^d -periodic function;

- this allows us to define the matrix

$$\mathbf{k}(t) := \int_0^t (\bar{a}(t') - \bar{a}) dt',$$

which by construction has zero-mean in time, i.e. $\int_0^1 \mathbf{k}(t) dt = 0$;

- integrating (2.4) with respect to space alone implies that $\langle \partial_t \chi_e \rangle = 0$, and thus for each $t \in \mathbb{R}$ we may define the *time corrector* $h_e(t, \cdot)$ as the the \mathbb{Z}^d -periodic $H^1(\mathbb{R}^d)$ solution of

$$-\Delta h_e(t, \cdot) = \partial_t \chi_e(t, \cdot), \quad \langle h_e(t, \cdot) \rangle = 0;$$

¹⁷The enhancement of diffusivity due to advection, from the point of view of homogenization theory, is by now classical, and we do not offer a complete summary of the relevant literature here. We instead refer the reader to [41, 54, 64] and the references therein.

- the purpose of introducing the space-homogenized matrix $\bar{\mathbf{a}}(t)$ and the time corrector $h_e(t, \cdot)$ is to ensure that for each $t \in \mathbb{R}$ the vector field

$$g_e(t, \cdot) := \mathbf{a}(t, \cdot)(e + \nabla \chi_e(t, \cdot)) - \bar{\mathbf{a}}(t)e + \nabla h_e(t, \cdot)$$

satisfies $\operatorname{div} g_e(t, \cdot) = 0$ and $\langle g_e(t, \cdot) \rangle = 0$; in turn, this allows us to write $g_e(t, \cdot)$ as the divergence of a skew-symmetric matrix $\mathbf{m}_e(t, \cdot)$, which is the zero-mean solution of

$$-\Delta \mathbf{m}_{e,ij}(t, \cdot) = \partial_i g_{e,j}(t, \cdot) - \partial_j g_{e,i}(t, \cdot).$$

With this notation, the parabolic two-scale expansion is given by

$$\tilde{\theta}_\varepsilon(t, x) := \theta(t, x) + \varepsilon \sum_{i=1}^d \partial_i \theta(t, x) \chi_{e_i} \left(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon} \right) + \varepsilon^2 \sum_{i,j=1}^d \mathbf{k}_{ij} \left(\frac{t}{\varepsilon^2} \right) \partial_{ij} \theta(t, x), \quad (2.7)$$

where θ is a solution of (2.3). The function θ is determined only once we specify a domain and a parabolic boundary condition.

In order to show that $\tilde{\theta}_\varepsilon$ is a good approximation of the solution θ_ε of (2.2) (again, by this we mean that one fixes a domain, and lets θ_ε match the values of θ on the corresponding parabolic boundary), we simply plug in the ansatz (2.7) into (2.2) and try to argue that the error we make is small. Using that θ solves (2.3) we arrive at:

$$\partial_t \tilde{\theta}_\varepsilon - \operatorname{div}(\mathbf{a}^\varepsilon \nabla \tilde{\theta}_\varepsilon) = f_\varepsilon, \quad (2.8)$$

where the “error” f_ε is given explicitly as

$$\begin{aligned} f_\varepsilon(t, x) := & \varepsilon \operatorname{div} \left(\sum_{i=1}^d (h_{e_i} \mathbf{I}_d - \chi_{e_i} \mathbf{a} - \mathbf{m}_{e_i}) \left(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon} \right) \nabla \partial_{x_i} \theta(t, x) \right) \\ & + \varepsilon \sum_{i=1}^d \left(\chi_{e_i} \left(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon} \right) \partial_{x_i} \partial_t \theta(t, x) + h_{e_i} \left(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon} \right) \partial_{x_i} \Delta \theta(t, x) \right) + \varepsilon^2 \sum_{i,j=1}^d \mathbf{k}_{ij} \left(\frac{t}{\varepsilon^2} \right) \partial_{ij} \partial_t \theta(t, x). \end{aligned} \quad (2.9)$$

The proof of the statement “the operator $\partial_t - \operatorname{div}(\mathbf{a}^\varepsilon \nabla)$ homogenizes as $\varepsilon \rightarrow 0$ to the operator $\partial_t - \operatorname{div}(\bar{\mathbf{a}} \nabla)$ ”, quantified by the convergence $\|\theta_\varepsilon - \theta\|_{L^2_{t,x}} \rightarrow 0$ as $\varepsilon \rightarrow 0$, now follows from (2.7), (2.8), and (2.9):

- (good) bounds on the solution χ_e of the cell problem (2.5) imply (good) bounds for the time correctors h_e , the skew-symmetric matrix \mathbf{m}_e , and the matrix \mathbf{k} ;
- coupled with (good) derivative estimates for the macroscopic function θ , which solves the homogenized problem (2.3), the aforementioned bounds show that the forcing f_ε appearing on the right side of (2.8) is $O(\varepsilon)$ (say in $L^2_t H_x^{-1}$);
- additionally, note that by the linearity of (2.2) we have that

$$\partial_t (\tilde{\theta}_\varepsilon - \theta_\varepsilon) - \operatorname{div}(\mathbf{a}^\varepsilon \nabla (\tilde{\theta}_\varepsilon - \theta_\varepsilon)) = f_\varepsilon, \quad (2.10)$$

and so by the standard parabolic energy estimate, $\tilde{\theta}_\varepsilon - \theta_\varepsilon$ is $O(\varepsilon)$ (say in $L^2_{t,x}$);

- Since θ_ε match the values of θ on the parabolic boundary of the spacetime domain, and $\tilde{\theta}_\varepsilon$ is an $O(\varepsilon)$ perturbation from θ , the difference $\tilde{\theta}_\varepsilon - \theta_\varepsilon$ is at most $O(\varepsilon)$ on the boundary (say in $L^\infty_{t,x}$);
- to conclude, note that from (2.7) and the above mentioned bounds we directly obtain that $\tilde{\theta}_\varepsilon - \theta$ is $O(\varepsilon)$ (say in $L^2_{t,x}$) by standard energy estimates.

The above items sweep under the rug the fact that the aforementioned bounds depend in a crucial way on the ellipticity constant of the matrix $\mathbf{a} = \kappa \mathbf{I}_d + \mathbf{s}$, on the ellipticity ratio of \mathbf{a} , which quantifies the size of the skew part \mathbf{s} to the size of the symmetric part $\kappa \mathbf{I}_d$. They also depend on the size of the derivatives of the macroscopic function θ , since these appear in the definition of f_ε in (2.9). These considerations are crucial for the quantitative homogenization estimates which we discuss next.

2.2. Quantitative homogenization for shear flows: Taylor diffusion. Stated as in §2.1, it appears that homogenization relies on *asymptotic scale separation* (the limit $\varepsilon \rightarrow 0$), which is a feature inconsistent with turbulence models.¹⁸ Asymptotic scale separation is also inconsistent with the construction of the vector field u in §3.1. There, ε represents the ratio of two consecutive length scales, which for a fixed parameter m in the iteration corresponds to $0 < \frac{\varepsilon_{m+1}}{\varepsilon_m} < 1$. Note however that for $m = 100$, say, the ratio $\frac{\varepsilon_{101}}{\varepsilon_{100}}$ is fixed, and cannot be sent to 0. This is why *quantitative homogenization* methods need, and should, be used! Rather than arguing that some quantities converge to others

¹⁸See e.g. the criticisms in [42, page 225] and [54, page 304].

in the limit of infinite scale separation, one needs to precisely quantify the length scales and time scales on which homogenization may be observed to have occurred, modulo small errors (which then need to be summed up across scales). Fortunately, quantitative homogenization rates for periodic coefficient fields are very explicit and quite well-understood (see [48, 65] for instance).

We will next highlight this quantitative homogenization perspective for the operator $\partial_t - \kappa\Delta + u \cdot \nabla$, when u is a *two-dimensional shear flow*. For parameters $a, \varepsilon > 0$, let us consider

$$u = u_{a,\varepsilon} = 2\pi a \varepsilon \cos\left(\frac{2\pi x_2}{\varepsilon}\right) e_1, \quad (2.11)$$

so that $0 < \varepsilon \ll 1$ is the scale of spatial oscillation (note that u is time independent), and $a \gg 1$ is proportional to the Lipschitz norm of u (note that $\|\nabla u\|_{L^\infty(\mathbb{T}^2)} = 4\pi^2 a$). Since we are two space dimensions, $\langle u \rangle = 0$, and $\operatorname{div} u = 0$, we may write

$$u_{a,\varepsilon} = \nabla^\perp \psi_{a,\varepsilon}, \quad \text{where} \quad \psi_{a,\varepsilon} = a\varepsilon^2 \sin\left(\frac{2\pi x_2}{\varepsilon}\right). \quad (2.12)$$

Associated to the stream function $\psi_{a,\varepsilon}$ we may associate the anti-symmetric matrix $\psi_{a,\varepsilon} \mathbf{I}_2^\perp$, where \mathbf{I}_2^\perp is the (rotation) matrix with components $(\mathbf{I}_2^\perp)_{11} = (\mathbf{I}_2^\perp)_{22} = 0$ and $(\mathbf{I}_2^\perp)_{21} = -(\mathbf{I}_2^\perp)_{12} = 1$. Then, in analogy to (2.2), we may rewrite

$$\partial_t - \kappa\Delta + u_{a,\varepsilon} \cdot \nabla = \partial_t - \operatorname{div}(\mathbf{a}^\varepsilon \nabla) \quad (2.13)$$

where for compatibility of notation we have defined

$$\mathbf{a}^\varepsilon(x) := \kappa \mathbf{I}_2 + \psi_{a,\varepsilon}(x) \mathbf{I}_2^\perp = \mathbf{a}\left(\frac{x}{\varepsilon}\right), \quad \text{with} \quad \mathbf{a}(y) := \kappa \mathbf{I}_2 + a\varepsilon^2 \sin(2\pi y_2) \mathbf{I}_2^\perp. \quad (2.14)$$

Here y is a placeholder for the fast space variable $\frac{x}{\varepsilon}$. In the example provided by (2.13)–(2.14) we may take advantage of the fact that \mathbf{a} is independent of time and of the first space coordinate to deduce that the correctors χ_e (which solve the cell problem (2.5)) are also independent of time and of the first space coordinates. In fact, we may quickly verify that the solution of (2.5) is given by

$$\chi_{e_1}(y) = \frac{a\varepsilon^2}{2\pi\kappa} \cos(2\pi y_2), \quad \text{and} \quad \chi_{e_2}(y) = 0, \quad (2.15)$$

which is clearly $\mathbb{Z} \times \mathbb{Z}^2$ -periodic. Plugging (2.15) into (2.4) we obtain

$$\bar{\mathbf{a}} = \kappa \mathbf{I}_2 + \begin{pmatrix} \frac{a^2 \varepsilon^4}{2\kappa} & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \kappa + \frac{1}{2}\kappa \cdot \left(\frac{a\varepsilon^2}{\kappa}\right)^2 & 0 \\ 0 & \kappa \end{pmatrix}. \quad (2.16)$$

The formula for the effective diffusion matrix $\bar{\mathbf{a}}$ given in (2.16) is called the *Taylor diffusion formula* [69] (specialized to the case of the shear flow in (2.11)); see also the work of Aris [3].

The Taylor diffusion formula for $\bar{\mathbf{a}}$ obtained in (2.16) is notable because $\bar{\mathbf{a}}$ is diagonal (note that \mathbf{a} had a non-constant anti-symmetric part), by definition it is constant in space and in time, and most notably, the enhancement of diffusion only occurs in the $e_1 \otimes e_1$ entry (a remnant of the fact that the u in (2.11) points in the e_1 direction but is independent of x_1). We also note that the quantity $\frac{a\varepsilon^2}{\kappa}$ which appears squared as a factor for the strength of the enhancement is dimensionless: a is the Lipschitz norm of a velocity field so it scales as $(\text{time})^{-1}$, the parameter ε has units of (length), while κ is a molecular diffusivity so it has units of $(\text{length})^2 \cdot (\text{time})^{-1}$.

At this point we have simply derived an explicit formula for the homogenized matrix $\bar{\mathbf{a}}$, but we have not said anything *quantitative* about the statement “the operator $\partial_t - \operatorname{div}(\mathbf{a}^\varepsilon \nabla)$ homogenizes as $\varepsilon \rightarrow 0$ to the operator $\partial_t - \operatorname{div}(\bar{\mathbf{a}} \nabla)$.” The following Lemma provides such a “sample” quantitative result; we do not claim that the following result is the sharpest possible, it is not the most general either; we present it here for pedagogical value, as it provides useful intuition about the quantitative bounds that may be extracted from the parabolic two-scale expansion.

Lemma 2.1 (Quantitative homogenization and Taylor diffusion). *Let \mathbf{a}^ε be defined by (2.14), and let $\bar{\mathbf{a}}$ be defined by (2.16). The operator $\partial_t - \operatorname{div}(\mathbf{a}^\varepsilon \nabla)$ homogenizes to the operator $\partial_t - \operatorname{div}(\bar{\mathbf{a}} \nabla)$, on*

$$\text{time-scales} \gtrsim a^{-1} \cdot \left(1 + \frac{a\varepsilon^2}{\kappa}\right) \quad \text{and} \quad \text{space-scales} \gtrsim \varepsilon \cdot \left(1 + \frac{a\varepsilon^2}{\kappa}\right). \quad (2.17)$$

More precisely:

- Define θ to be the \mathbb{T}^2 -periodic solution of the Cauchy problem for $\partial_t \theta - \operatorname{div}(\bar{\mathbf{a}} \nabla \theta) = 0$ (cf. (2.3)), with smooth initial datum $\theta|_{t=0} = \theta_{\text{in}}$.
- For $R > 0$ define the parabolic cylinders adapted to $\bar{\mathbf{a}}$, as $Q_R = Q_R(t_0, x_0) = (t_0, x_0) + (-\frac{R^2}{\kappa}, 0] \times E_R(0)$, where E_R is the ellipsoid defined by $\{x: \kappa^{\frac{1}{2}} |\bar{\mathbf{a}}^{-\frac{1}{2}} x| \leq R\} = \kappa^{-\frac{1}{2}} \bar{\mathbf{a}}^{\frac{1}{2}} B_R(0)$. In our case, due to (2.16), this ellipsoid has axes of length $\sim R(1 + \frac{a\varepsilon^2}{\kappa})$ in the e_1 direction, and $\sim R$ in the e_2 direction.
- For every such parabolic cylinder Q_R with $Q_{2R} \subseteq [0, 1] \times \mathbb{T}^2$, let θ_ε be the solution of $\partial_t \theta_\varepsilon - \operatorname{div}(\mathbf{a}^\varepsilon \nabla \theta_\varepsilon) = 0$ in Q_R , with (parabolic) boundary data $\theta_\varepsilon \equiv \theta$ on $\partial_{\text{par}} Q_R$.

- Then, we have that

$$\|\theta_\varepsilon - \theta\|_{L_t^\infty \underline{L}_x^2(Q_R)} \lesssim (1 + \frac{a\varepsilon^2}{\kappa}) \frac{\varepsilon}{R} \|\theta\|_{\underline{L}_{x,t}^2(Q_{2R})}. \quad (2.18)$$

Here, $L_t^\infty \underline{L}_x^2(Q_R)$ and $\underline{L}_{x,t}^2(Q_R)$ denote the usual volume-normalized norm (e.g. integrals \int_{Q_R} are replaced by \int_{Q_R}). In particular, when $R \gg \varepsilon(1 + \frac{a\varepsilon^2}{\kappa})$, we see that $\|\theta_\varepsilon - \theta\|_{L_t^\infty \underline{L}_x^2(Q_R)} \ll \|\theta\|_{\underline{L}_{x,t}^2(Q_{2R})}$, which matches the heuristic claim regarding space-scales in (2.17).

Brief proof of Lemma 2.1. We wish to appeal to the parabolic two-scale expansion described in §2.1. We have already computed the correctors with slope e , namely χ_e , in (2.15). Since $\mathbf{a} = \mathbf{a}(y_2)$ and $\chi_e = \chi_e(y_2)$ are both independent of t , the space homogenized matrix $\bar{\mathbf{a}}(t)$ is time independent, and is equal to $\bar{\mathbf{a}}$ for all $t \in \mathbb{R}$. Hence, the matrix $\mathbf{k}(t) = 0$ for all $t \in \mathbb{R}$, and the time correctors also vanish identically, $h_e(t, \cdot) = 0$ for all $t \in \mathbb{R}$. Then, it follows that $g_{e_1}(t, y) = -\frac{a^2 \varepsilon^4}{2\kappa} \cos(4\pi y_2) e_1$ and $g_{e_2}(t, y) = -a\varepsilon^2 \sin(2\pi y_2) e_1$, so that the skew-symmetric matrices $\mathbf{m}_e(t, \cdot)$ are given explicitly by $\mathbf{m}_{e_1}(t, y) = \frac{a^2 \varepsilon^4}{8\pi\kappa} \sin(4\pi y_2) \mathbf{I}_2^\perp$ and $\mathbf{m}_{e_2}(t, y) = -\frac{a\varepsilon^2}{2\pi} \cos(2\pi y_2) \mathbf{I}_2^\perp$.

Using the explicit computations in the previous paragraph, the equation (2.3) for θ may also be written as

$$\partial_t \theta - \kappa \left(1 + \frac{a^2 \varepsilon^4}{2\kappa^2}\right) \partial_{11} \theta - \kappa \partial_{22} \theta = 0. \quad (2.19)$$

Moreover, the parabolic two-scale expansion is then given by (2.7) as

$$\tilde{\theta}_\varepsilon(t, x) := \theta(t, x) + \varepsilon \partial_1 \theta(t, x) \chi_{e_1}\left(\frac{x_2}{\varepsilon}\right) = \theta(t, x) + \frac{a\varepsilon^3}{2\pi\kappa} \cos\left(\frac{2\pi x_2}{\varepsilon}\right) \partial_1 \theta(t, x). \quad (2.20)$$

Additionally, the “error” term f_ε defined in (2.9) may be computed explicitly as

$$\begin{aligned} f_\varepsilon(t, x) &= -\varepsilon \operatorname{div} \left((\chi_{e_1} \mathbf{a} + \mathbf{m}_{e_1}) \left(\frac{x_2}{\varepsilon}\right) \nabla \partial_1 \theta(t, x) + \mathbf{m}_{e_2} \left(\frac{x_2}{\varepsilon}\right) \nabla \partial_2 \theta(t, x) \right) + \varepsilon \chi_{e_1}\left(\frac{x_2}{\varepsilon}\right) \partial_1 \partial_t \theta(t, x) \\ &= -\varepsilon \operatorname{div} \left(\frac{a\varepsilon^2}{2\pi} \cos\left(\frac{2\pi x_2}{\varepsilon}\right) \nabla \partial_1 \theta(t, x) - \frac{a^2 \varepsilon^4}{8\pi\kappa} \sin\left(\frac{4\pi x_2}{\varepsilon}\right) \nabla^\perp \partial_1 \theta(t, x) \right) \\ &\quad + \varepsilon \operatorname{div} \left(\frac{a\varepsilon^2}{2\pi} \cos\left(\frac{2\pi x_2}{\varepsilon}\right) \nabla^\perp \partial_2 \theta(t, x) \right) + \varepsilon \partial_1 \left(\frac{a\varepsilon^2}{2\pi\kappa} \cos\left(\frac{2\pi x_2}{\varepsilon}\right) \partial_t \theta(t, x) \right) \\ &=: \operatorname{div} v_\varepsilon(t, x), \end{aligned} \quad (2.21)$$

for an implicitly defined vector v_ε , and where we recall that $\nabla^\perp = \mathbf{I}_2^\perp \nabla = (-\partial_2, \partial_1)$.

In order to conclude, we at last perform some PDE estimates. First, we need some standard bounds for θ which are just standard pointwise estimates for the heat equation after a change of variables. These say that, for every $\ell \in \{0, 1, 2, \dots\}$,

$$(\kappa^{-\frac{1}{2}} R)^\ell \|(\bar{\mathbf{a}}^{\frac{1}{2}} \nabla)^\ell \theta\|_{L_{x,t}^\infty(Q_R)} + (\kappa^{-\frac{1}{2}} R)^{2\ell} \|\partial_t^\ell \theta\|_{L_{x,t}^\infty(Q_R)} \leq C_\ell \|\theta\|_{\underline{L}_{x,t}^2(Q_{2R})}. \quad (2.22)$$

Next, we record a bound which is immediate from (2.21) and standard global energy estimates for (2.19) namely

$$\begin{aligned} \|v_\varepsilon\|_{L_{x,t}^2(Q_R)} &\lesssim (a\varepsilon^3 + \frac{a^2 \varepsilon^5}{\kappa}) \|\partial_1 \nabla \theta\|_{L_{x,t}^2(Q_R)} + a\varepsilon^3 \|\partial_{22} \theta\|_{L_{x,t}^2(Q_R)} + \frac{a\varepsilon^3}{\kappa} \|\partial_t \theta\|_{L_{x,t}^2(Q_R)} \\ &\lesssim \frac{\kappa\varepsilon}{R^2} \left(1 + \frac{a\varepsilon^2}{\kappa}\right) \|\theta\|_{\underline{L}_{x,t}^2(Q_{2R})}, \end{aligned} \quad (2.23)$$

where the implicit constant in the \lesssim symbol is universal. With this bound, we return to (2.10) and recall the parabolic problem

$$\partial_t (\tilde{\theta}_\varepsilon - \theta_\varepsilon) - \operatorname{div}(\mathbf{a}^\varepsilon \nabla (\tilde{\theta}_\varepsilon - \theta_\varepsilon)) = \operatorname{div} v_\varepsilon. \quad (2.24)$$

Note that (2.20) gives $(\tilde{\theta}_\varepsilon - \theta_\varepsilon)(t, x) = \frac{a\varepsilon^3}{2\pi\kappa} \cos\left(\frac{2\pi x_2}{\varepsilon}\right) \partial_1 \theta(t, x)$ for $(t, x) \in \partial_{\text{par}} Q_R$, and so $\tilde{\theta}_\varepsilon - \theta_\varepsilon$ does not vanish identically on $\partial_{\text{par}} Q_R$; instead we have

$$\|\tilde{\theta}_\varepsilon - \theta_\varepsilon\|_{L_{x,t}^\infty(\partial_{\text{par}} Q_R)} \lesssim \frac{a\varepsilon^3}{\kappa} \|\partial_1 \theta\|_{L^\infty(Q_R)} \leq \frac{\varepsilon}{R} \|\theta\|_{\underline{L}_{x,t}^2(Q_{2R})}.$$

The right side is bounded by that of (2.18). Therefore, by the maximum principle, we can correct this discrepancy at the boundary (and just ignore it for our purposes).

We continue therefore by pretending that $\tilde{\theta}_\varepsilon - \theta_\varepsilon$ vanishes on $\partial_{\text{par}} Q_R$, and return to (2.24) to perform a standard energy estimate. We obtain:

$$\|\tilde{\theta}_\varepsilon - \theta_\varepsilon\|_{L_t^\infty \underline{L}_x^2(Q_R)}^2 + \kappa \|\nabla(\tilde{\theta}_\varepsilon - \theta_\varepsilon)\|_{L_{x,t}^2(Q_R)}^2 \lesssim \kappa^{-1} \|v_\varepsilon\|_{L_{x,t}^2(Q_R)}^2.$$

We continue by ignoring the gradient term on the left side and appeal to (2.20), (2.22) and (2.23) to obtain

$$\begin{aligned} \|\tilde{\theta}_\varepsilon - \theta\|_{L_t^\infty \underline{L}_x^2(Q_R)} &\leq \|\tilde{\theta}_\varepsilon - \theta_\varepsilon\|_{L_t^\infty \underline{L}_x^2(Q_R)} + \|\theta_\varepsilon - \theta\|_{L_t^\infty \underline{L}_x^2(Q_R)} \\ &\lesssim \kappa^{-\frac{1}{2}} (\kappa^{-1} R^2)^{\frac{1}{2}} \|v_\varepsilon\|_{L_{x,t}^2(Q_R)} + \frac{\varepsilon}{R} \|\theta\|_{L^2(Q_{2R})} \lesssim (1 + \frac{a\varepsilon^2}{\kappa}) \frac{\varepsilon}{R} \|\theta\|_{\underline{L}_{x,t}^2(Q_{2R})}. \end{aligned}$$

This concludes the proof of (2.18). \square

3. AN OUTLINE OF THE PROOF OF THEOREM 1.2

The proof consists of two parts: the construction of multi-scale the vector field u (see §3.1), and the analysis of the passive scalars θ^K via quantitative homogenization (see §3.2–§3.4). Throughout this section, it is instructive to keep in mind the intuition gained from the discussion in §2.

3.1. Construction of the vector field u . The proof of Theorem 1.2 is the most difficult in the lowest nontrivial dimension, namely $d = 2$, for the same geometric reasons that the Onsager conjecture was hardest to establish in the two-dimensional case [45]. We present the $d = 2$ construction next.

The advantage of the case $d = 2$ is notational: zero-mean divergence-free vector fields u are given as the rotated gradient of a scalar stream function ϕ (with associated anti-symmetric matrix $I_2^\perp \phi$); as such, we avoid the notational care one must use when handling vectors, or when working with the curl operator.

For $m \geq 0$, we recursively construct a sequence of periodic, in $(t, x) \in [0, 1] \times \mathbb{T}^2$, $C_{x,t}^\infty$ smooth vector fields $u_m = \nabla^\perp \phi_m$, and then define $u = \lim_{m \rightarrow \infty} u_m$; here the limit is taken with respect to the $C_{x,t}^\alpha$ topology. We initialize the construction with $\phi_0 = 0$, so that $u_0 = 0$. For any $m \geq 1$, having constructed the smooth, periodic, incompressible vector field u_{m-1} with associated stream function ϕ_{m-1} , the induction step $m-1 \mapsto m$ consists in the construction of a suitable stream function “increment” $\phi_m - \phi_{m-1}$. For this purpose, we make the following initial parameter choices:

- Let $\{\varepsilon_m\}_{m \geq 0}$ be a decreasing sequence of scales, with $\varepsilon_m \rightarrow 0$ as $m \rightarrow \infty$. We set $\varepsilon_0 = 1$, and $\varepsilon_1 = \Lambda^{-1}$ for a large integer $\Lambda \gg 1$, chosen at the end of the proof. At each step $m \geq 1$, we view ε_m as the “fast scale” of oscillation, and ε_{m-1} as the “slow scale”.¹⁹ Based on previous experience with multi-scale constructions in turbulence (see e.g. [29]), we expect that it will be technically (very!) convenient to consider the sequence ε_m to decay “slightly super-exponentially”, meaning that

$$\varepsilon_m \simeq \varepsilon_{m-1}^q, \quad \text{for some } 0 < q - 1 \ll 1, \quad (3.1)$$

for all $m \geq 2$. Here, the “ \simeq ” is not an “=” sign because we need ε_m^{-1} to be an integer for all $m \geq 0$ (in order to enforce \mathbb{T}^2 -periodicity). How small the positive parameter $q - 1$ needs to be taken is chosen as part of the proof.

- We let $a_m \gg 1$ denote the Lipschitz norm (up to universal constants) of the stream function increment $\phi_m - \phi_{m-1}$. If we think of the increment $\phi_m - \phi_{m-1}$ as oscillating at frequency ε_m , then in order to ensure uniform boundedness of the velocity fields $\{u_m\}_{m \geq 0}$ in $C_t^0 C_x^\alpha$, it is natural and necessary to define

$$a_m := \varepsilon_m^{\alpha-1}, \quad \text{for all } m \geq 0. \quad (3.2)$$

Note that since $\alpha < 1$, we have $a_m \rightarrow \infty$ as $m \rightarrow \infty$.

3.1.1. A naive attempt for all $0 < \alpha < 1$. With $\{\varepsilon_m\}_{m \geq 0}$ and $\{a_m\}_{m \geq 0}$ fixed according to (3.1)–(3.2), and keeping in mind the result of Lemma 2.1, a “naive” attempt at defining the stream function increment $\phi_m - \phi_{m-1}$ is as follows.

Based on (2.12) we are tempted to declare that the stream function increment is given by $a_m \varepsilon_m^2 \sin(\frac{2\pi x_2}{\varepsilon_m})$. Note however that in this case the homogenized matrix (2.16) presents an enhancement only in the $e_1 \otimes e_1$ component, which is undesirable. The natural fix to this issue is to alternate the direction of shearing, by considering a stream function which sequentially in time alternates the stream functions $a_m \varepsilon_m^2 \sin(\frac{2\pi x_2}{\varepsilon_m})$ and $a_m \varepsilon_m^2 \sin(\frac{2\pi x_1}{\varepsilon_m})$, “glued” by a partition of unity in time, with characteristic time-scale proportional to some $0 < \tau_m \ll 1$, to be determined. We would like however these shears acting on different directions to not interact/interfere with each other (see also [16, 27, 47]), and hence we leave room in which no shearing occurs, whenever we switch directions. As such, we define

$$\Psi_k(x) := \begin{cases} \sin(2\pi x_1), & k \equiv 1 \pmod{4}, \\ \sin(2\pi x_2), & k \equiv 3 \pmod{4}, \\ 0, & k \equiv 0 \pmod{2}, \end{cases}$$

and we attempt to declare

$$\phi_m(t, x) - \phi_{m-1}(t, x) := \sum_{k \in \mathbb{Z}} \zeta\left(\frac{t}{\tau_m} - k\right) a_m \varepsilon_m^2 \Psi_k\left(\frac{x}{\varepsilon_m}\right), \quad (3.3)$$

where $1_{[-\frac{1}{3}, \frac{1}{3}]} \leq \zeta \leq 1_{[-\frac{2}{3}, \frac{2}{3}]}$ is a C^∞ smooth unit-scale bump function with $\int_{\mathbb{R}} \zeta^2(t) dt = 1$, and which induces a partition of unity via $\sum_{k \in \mathbb{Z}} \zeta(\cdot - k) \equiv 1$.

¹⁹The discussion in §2.2 refers to the fast scale ε and the slow scale 1. As such, in this new context it is the ratio $\frac{\varepsilon_m}{\varepsilon_{m-1}} < 1$ which plays the role of the small parameter ε in §2.2.

By construction, the stream function increment in (3.3) is $4\tau_m$ -periodic in time, and since the formula (2.4) for the homogenized matrix involves averaging in time, we expect that enhancement will occur in both the $e_1 \otimes e_1$ and $e_2 \otimes e_2$ components, as desired, with the “amount of enhancement” being 25% of the one in (2.16). Note however that in order to see this enhancement, Lemma 2.1 dictates that we “zoom out” in both time and space; in turn, this imposes restrictions on τ_m and ε_m .

The first condition in (2.17) dictates a constraint on the time scale τ_m , which determines the periodicity in time (inverse frequency) of the stream function increment in (3.3). We need the amount of time that each of the stream functions Ψ_k are active to be long enough to allow for homogenization to be observed; this dictates:

$$\tau_m \gg a_m^{-1} \cdot \left(1 + \frac{a_m \varepsilon_m^2}{\kappa_m}\right) = a_m^{-1} + \frac{\varepsilon_m^2}{\kappa_m} \geq \frac{\varepsilon_m^2}{\kappa_m}. \quad (3.4)$$

The second condition in (2.17) dictates that the length scale corresponding to “slow oscillations” (meaning, the scale of oscillation of the “slow” stream function ϕ_{m-1}), is large enough for homogenization to be observed; this dictates:

$$\varepsilon_{m-1} \gg \varepsilon_m \cdot \left(1 + \frac{a_m \varepsilon_m^2}{\kappa_m}\right) \geq \frac{a_m \varepsilon_m^3}{\kappa_m}. \quad (3.5)$$

With (3.4)–(3.5), Lemma 2.1 indicates that upon “zooming out” from scale ε_m to scale ε_{m-1} ,²⁰ the enhancement of diffusion is given by $\kappa I_2 \mapsto \kappa I_2 + \frac{1}{8}\kappa \cdot \left(\frac{a_m \varepsilon_m^2}{\kappa}\right)^2 I_2$; the division by 4 when compared to (2.16) is due to time-averaging.

If the above-described scenario can be justified rigorously, iteratively in m , then we have set up the parameters $\{\varepsilon_m\}_{m \geq 0}$, $\{a_m\}_{m \geq 0}$, and $\{\tau_m\}_{m \geq 0}$, such that when “coarse-graining” at scale ε_m ²¹ the advection-diffusion equation (1.1), with vector field $u = \nabla^\perp \phi$, with ϕ defined by summing over $m \geq 0$ the increments in (3.3), experiences an effective diffusivity κ_m (see §3.2), which satisfies the recursion relation

$$\kappa_{m-1} := \kappa_m + \frac{1}{8}\kappa_m \cdot \left(\frac{a_m \varepsilon_m^2}{\kappa_m}\right)^2. \quad (3.6)$$

The recursion (3.6) is to be “initialized” with $\kappa_M = \kappa$, for an integer $M \gg 1$ which is to be determined such that ε_M is proportional to the dissipation length-scale associated with the advection-diffusion equation (1.1) with a $C_{x,t}^\alpha$ vector field u . With this initialization, the recursion (3.6) is to be solved backwards: $M \rightarrow M-1 \rightarrow \dots \rightarrow m \rightarrow m-1 \rightarrow \dots \rightarrow 0$, with the hope that $\kappa = \kappa_M < \kappa_{M-1} < \dots < \kappa_m < \kappa_{m-1} < \dots < \kappa_0$, and that $\kappa_0 = O(1)$ as $\kappa \rightarrow 0$.

Let us next inquire whether the parameter choices (3.1), (3.2), together with the recursion (3.6), are consistent with the constraints (3.4), (3.5), and with the fact that $\kappa_0 = O(1)$. Assuming for simplicity that $\kappa = \kappa_M = \varepsilon_M^{1+\alpha+\gamma}$ for some $M \geq 1$ ²²; then, it is not hard to verify that the relation

$$\kappa_m \simeq \varepsilon_m^{1+\alpha+\gamma}, \quad \text{where} \quad \gamma = (q-1) \frac{1+\alpha}{q+1}, \quad (3.7)$$

for all $M \geq m \geq 0$, is consistent with (3.6), where we recall that $0 < q-1 \ll 1$. Moreover, setting $m=0$ in (3.7) yields $\kappa_0 \simeq 1$, as desired, independently of the value of $M \gg 1$ (and hence of $\kappa \ll 1$).

The constraint (3.4) is then used as to pick a time-scale that satisfies

$$\tau_m \gg a_m^{-1} + \varepsilon_m^2 \kappa_m^{-1} = \varepsilon_m^{1-\alpha} + \varepsilon_m^{1-\alpha-\gamma} \simeq \varepsilon_m^{1-\alpha-\gamma}. \quad (3.8)$$

Such a choice is permissible if $\tau_m \leq 1^{23}$ for all $m \geq 0$, i.e., as soon as $1 - \alpha - \gamma > 0$. In light of the definition of γ in (3.7), this is indeed possible for all values $\alpha \in (0, 1)$, by letting $1 < q < \frac{1}{\alpha}$. Note that as $\alpha \rightarrow 1^-$, this implies $q \rightarrow 1^+$.

3.1.2. What goes wrong with this naive construction? The discussion in §3.1.1 indicates that for any $\alpha \in (0, 1)$ we may construct a $C_{x,t}^\alpha$ vector field (with stream function obtained by summing (3.3) in $m \geq 1$), such that the sequence of renormalized diffusivities are compatible with the recursion relation (3.6), and the space-time separation is compatible with (3.4) and (3.5), as dictated by Lemma 2.1; additionally, the amount of diffusion experienced by the passive scalar when “coarse-grained” at $O(1)$ scales is itself $O(1)$, independently of the value of the molecular diffusion $\kappa \ll 1$.

In discussing the above heuristic, *a subtle (and fatal) point was overlooked*: the homogenization of a shear flow discussed in Lemma 2.1 assumed that the shear flow $u = u_{a,\varepsilon}$ had zero-mean when averaged over one periodic cell of

²⁰It is this $m \mapsto m-1$ homogenization step which goes from small scales to larger ones, and this is the motivation for referring to the cascade of effective diffusivities as being “backwards”.

²¹See [4,6] for a general theory of *quantitative theory of coarse-graining* for elliptic and parabolic PDEs with multi-scale oscillating coefficients.

²²That is, we are letting $\kappa = \kappa_M \rightarrow 0$ along the designated sequence $\varepsilon_M^{1+\alpha+\gamma} \rightarrow 0$, as $M \rightarrow \infty$.

²³The choice of a_m in (3.2), together with the definition (3.3) implies the uniform-in- m space regularity estimate of the velocity increments: $\|u_m - u_{m-1}\|_{C_t^\alpha C_x^0} = \|\nabla^\perp(\phi_m - \phi_{m-1})\|_{C_t^\alpha C_x^0} \simeq a_m \varepsilon_m^{1-\alpha} \simeq 1$. On the other hand, if we wish to estimate uniform-in- m time regularity of the velocity increments, we see that $\|u_m - u_{m-1}\|_{C_t^\alpha C_x^0} = \|\nabla^\perp(\phi_m - \phi_{m-1})\|_{C_t^\alpha C_x^0} \simeq \tau_m^{-\alpha} a_m \varepsilon_m \simeq \tau_m^{-\alpha} \varepsilon_m^\alpha \ll \varepsilon_m^{\alpha(\alpha+\gamma)}$, where in the last inequality we have used (3.8). For any $\alpha > 0$ and $q > 1$ we thus obtain C_x^α -regularity of our vector field for free.

the fast variable $y = \frac{x}{\varepsilon}$. Indeed, the $\cos(2\pi y_2)$ present in the definition of $u_{a,\varepsilon}$ (see (2.11)) ensures that this velocity field has zero-mean over $\mathbb{T}^2(dy)$. It was the very fact that $\langle u_{a,\varepsilon} \rangle = 0$ that allowed us to write this velocity field as the rotated gradient of a periodic stream function $\psi_{a,\varepsilon}$ (see (2.12)), which was at the core of our homogenization analysis.

Returning to the construction in §3.1.1, for each $m \geq 1$ we note that the velocity field u_m *does not have zero-mean when averaged over a periodic cell of the fast variable* $\frac{x}{\varepsilon_m}$. Indeed, the “macroscopic” velocity field u_{m-1} is *essentially a constant* from the point of view of the “microscopic” velocity increment $u_m - u_{m-1}$, which is $(\varepsilon_m \mathbb{T})^2$ -periodic. To be precise, recursion (3.3) shows that we have $\int_{(\varepsilon_m \mathbb{T})^2} (u_m - u_{m-1})(t, \cdot) = 0$ for all $t \in \mathbb{R}$, while for any $1 \leq j \leq m-1$, we have²⁴ $\int_{(\varepsilon_m \mathbb{T})^2} (u_{m-j} - u_{m-j-1})(t, \cdot) \simeq \varepsilon_{m-j}^\alpha \neq 0$, for $t \in \cup_{k \in 2\mathbb{Z}+1} [\tau_{m-j}(k - \frac{1}{3}), \tau_{m-j}(k + \frac{1}{3})]$. Note that $\varepsilon_{m-j}^\alpha \gg \varepsilon_m^\alpha$ for any $1 \leq j \leq m-1$, so not only is the macroscopic velocity field u_{m-1} essentially a constant from the point of view of microscopic oscillations, in relative terms it is a *large constant*.

The issue is that presence of a large constant drift velocity, superimposed on a zero-mean shear flow, is able to remove most of the enhancement of the diffusivity; this is so because generically this constant drift will point in an irrational/Diophantine direction, causing *averaging to overtake homogenization*. This is explained quantitatively in [9, Appendix A]. In the physics literature this phenomenon—that a non-constant mean flow interferes and potentially destroys Taylor diffusion—is termed the “sweeping effect” [12, 38, 54]. To sum up, the construction presented in §3.1.1 does not work, at least not “as is”.

3.1.3. Small-scale features are transported by the large-scale flow: $\alpha < \frac{1}{3}$. To overcome the issue raised at the end of §3.1.2 we recall one of the cornerstones of phenomenological turbulence theories, the observation that microscopic “eddies” are transported by the surrounding macroscopic flow, for a suitable amount of time [42].

Roughly speaking, the idea put forth in [9]²⁵ is that the microscopic/fast shear flows should *not* be naively superimposed on top of the existing macroscopic/slow shear flows (as (3.3) dictates); *instead* the microscopic/fast shear flows should be added in a frame of reference that moves according to the macroscopic/slow velocity field (which is to say, in the Lagrangian coordinates of this slow velocity), so that in this frame the shear flows continue to have zero mean (to leading order).

For this purpose, for each $[0, 1] \times \mathbb{T}^2$ -periodic, divergence-free velocity field u_m , we define the associated flow maps $X_m = X_m(t, x; s)$ by solving the ODE

$$\frac{d}{dt} X_m(t, x; s) = u_m(t, X_m(t, x; s)), \quad X_m(s, x; s) = x,$$

where $t, s \in [0, 1]$ and $x \in \mathbb{T}^2$. Here we use the notation introduced in §3.1; in particular, for each $m \geq 1$ the velocity fields $u_m \in C_{x,t}^\infty$. We shall denote by $X_m^{-1} = X_m^{-1}(x, t; s)$ the corresponding inverse maps (sometimes referred to as the “back-to-labels map”), satisfying the relations $X_m^{-1}(t, X_m(t, x; s); s) = x$ and $X_m(t, X_m^{-1}(x, t; s); s) = x$.²⁶ Classical ODE theory dictates that the Lagrangian maps $x \mapsto X_m(\cdot, x; \cdot)$ and their inverse flows $x \mapsto X_m^{-1}(\cdot, x; \cdot)$ remain close to the identity map, in the sense that $|\nabla X_m(x, t; s) - I_2| + |\nabla X_m^{-1}(x, t; s) - I_2| \ll 1$, only for time intervals which are small multiples of the Courant-Friedrichs-Lax time, namely for $|t - s| \ll \|\nabla u_m\|_{L_{x,t}^\infty}^{-1}$. Thus, the usage of Lagrangian coordinates necessitates the introduction of a new time-scale $\{\tau_m''\}_{m \geq 1}$, which needs to satisfy the relation²⁷

$$\tau_m'' \|\nabla u_m\|_{L_{x,t}^\infty} \ll 1 \quad \text{or, equivalently,} \quad \tau_m'' a_m = \tau_m'' \varepsilon_m^{\alpha-1} \ll 1. \quad (3.9)$$

According to this new time-scale, for every $\ell \in \mathbb{Z}$ we denote²⁸

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

the forward and backward flows of u_m , initialized at time $\ell \tau_m''$.

²⁴Here we use $\int_0^{\varepsilon_m} \cos(\frac{2\pi x_1}{\varepsilon_{m-j}}) dx_1 = \varepsilon_{m-j}^{\alpha+1} (2\pi \varepsilon_m)^{-1} \sin(\frac{2\pi \varepsilon_m}{\varepsilon_{m-j}}) \simeq \varepsilon_{m-j}^\alpha$ since $\frac{\varepsilon_m}{\varepsilon_{m-j}} \ll 1$.

²⁵In the context of the Onsager conjecture, this idea (a proto-version of it) was introduced first in [29], and played a major role in all subsequent works, leading to [16, 44, 45, 47, 60]. In the context of homogenization of advection-diffusion equations, a version of this idea appeared in [56].

²⁶Note that the maps $x \mapsto X_m(t, x; s)$ and $x \mapsto X_m^{-1}(t, x; s)$ are volume-preserving (since $\operatorname{div} u_m = 0$), while the maps $x \mapsto X_m(t, x; s) - x$ and $x \mapsto X_m^{-1}(t, x; s) - x$ are $[0, 1] \times \mathbb{T}^2$ -periodic (since so is u_m).

²⁷We note here a difference with the notation in [9]: there, τ_m'' is chosen to be small with respect to the CFL time-scale of u_{m-1} , i.e. $\tau_m'' a_{m-1} \ll 1$. In this note we choose the slightly more intuitive notation, which relates τ_m'' to the CFL time-scale of u_m , i.e. $\tau_m'' a_m \ll 1$; see (3.9).

²⁸As before, we have a difference with the notation with respect to [9]: here we write $X_m(\cdot, \cdot; \ell \tau_m'')$, whereas in [9, Equation (2.36)] we wrote $X_{m-1}(\cdot, \cdot; \ell \tau_m'')$.

With the above notation, for $m \geq 1$ we define the stream function increment by²⁹

$$\phi_m(t, x) - \phi_{m-1}(t, x) := \sum_{k, \ell \in \mathbb{Z}} \zeta\left(\frac{t}{\tau_m} - k\right) \widehat{\zeta}\left(\frac{t}{\tau_{m-1}''} - \ell\right) a_m \varepsilon_m^2 \Psi_k\left(\frac{X_{m-1, \ell}^{-1}(t, x)}{\varepsilon_m}\right), \quad (3.10)$$

where $\widehat{\zeta}$ is another C^∞ smooth unit-scale bump function, which smoothly approximates the cutoff function $1_{[-\frac{1}{2}, \frac{1}{2}]}$, and such that its integer shifts nearly form a partition of unity: $\sum_{\ell \in \mathbb{Z}} \widehat{\zeta}(\cdot - \ell) \simeq 1$.³⁰ Note that the cutoff functions $\zeta(\cdot - k)$, the parameters $\varepsilon_m, a_m, \tau_m$, and also the alternating shears Ψ_k appearing in (3.10) are the same as in (3.3). At this stage we only assume that the time-scale τ_{m-1}'' satisfies the smallness condition from (3.9), but as we shall see next, a largeness constraint emerges (see (3.11) below). With the new definition of the stream function increment in (3.10) replacing the previous naive attempt from (3.3), the velocity increment is defined again as as $u_m - u_{m-1} = \nabla^\perp(\phi_m - \phi_{m-1})$.

The usefulness of (3.10) is that on the support of each cutoff function $\widehat{\zeta}(\frac{t}{\tau_{m-1}''} - \ell)$, the “Lagrangian increment” $(\phi_m - \phi_{m-1})(\cdot, X_{m-1, \ell}(\cdot, x))$ behaves as $a_m \varepsilon_m^2 \Psi_k(\frac{x}{\varepsilon_m})$, which has zero-mean with respect to the fast space variable $\frac{x}{\varepsilon_m}$. This overcomes the issue raised at the end of §3.1.2. However, in order to use the same heuristics as in §3.1.1, leading to the recursion relation (3.6) and asymptotic (3.7) for the cascade of renormalized diffusivities, we need to ensure that the Lagrangian flow maps $X_{m-1, \ell}^{-1}(\cdot, \cdot)$ and the new time cutoffs $\widehat{\zeta}(\frac{t}{\tau_{m-1}''} - \ell)$ appearing in (3.10), do not interfere with the quantitative homogenization picture painted in §3.1.1. This necessitates that the new periodic time oscillation which we have introduced (on time-scale τ_{m-1}'') *does not interfere*³¹ with the homogenization problem which produced (3.6); put differently, this necessitates that from the point of view of a function oscillating at time-scale τ_m , functions varying on time-scale τ_{m-1}'' are essentially a constant, resulting in the new parameter constraint

$$\tau_{m-1}'' \gg \tau_m. \quad (3.11)$$

Together, the constraint on the new time parameter τ_m'' are

$$\tau_m \ll \tau_{m-1}'' \ll a_{m-1}^{-1}. \quad (3.12)$$

We also recall that τ_m needs to be chosen to satisfy (3.4).

At last, we try to see if we can choose the parameters in order satisfy all of the above-mentioned constraints. The two inequalities in (3.12) indicate that a permissible choice for τ_{m-1}'' can be made if and only if τ_m is chosen to satisfy $a_{m-1}^{-1} \gg \tau_m$. In turn, the lower bound for τ_m established previously in (3.8) shows that such a permissible choice for τ_m is possible if and only if $a_{m-1}^{-1} \gg \varepsilon_m^{1-\alpha-\gamma}$, where we recall that γ is defined by (3.7). Recalling the definition of a_{m-1} , and the ε_m to ε_{m-1} relation from (3.1), we arrive at the parameter constraint

$$\varepsilon_m^{\frac{1-\alpha}{q}} \simeq \varepsilon_{m-1}^{1-\alpha} = a_{m-1}^{-1} \gg \varepsilon_m^{1-\alpha-\gamma}. \quad (3.13)$$

Inequality (3.13) may be satisfied for all $m \geq 1$ if $\varepsilon_1 = \Lambda^{-1}$ is chosen to be sufficiently small (to absorb the implicit constant in the \gg symbol), and if $\frac{1-\alpha}{q} < 1 - \alpha - \gamma$. Recalling that $\gamma = (q-1)\frac{1+\alpha}{q+1}$, that $\alpha \in (0, 1)$ and $q > 1$, (3.13) thus necessitates:

$$\frac{1-\alpha}{q} < 1 - \alpha - \gamma \Leftrightarrow q\gamma < (q-1)(1-\alpha) \Leftrightarrow \frac{q(1+\alpha)}{q+1} < 1 - \alpha \Leftrightarrow \alpha < \frac{1}{2q+1}.$$

Since $q > 1$, we arrive at the *Onsager-supercriticality constraint* $\alpha < \frac{1}{3}$ present in the statement of Theorem 1.2.

3.1.4. The construction of u : order of choosing parameters. To sum up, for a given $\alpha \in (0, \frac{1}{3})$, we first choose $q > 1$ so that $(2q+1)\alpha < 1$. Then, we define $\{\varepsilon_m\}_{m \geq 0}$ according to (3.1), and $\{a_m\}_{m \geq 0}$ according to (3.2). In light of the lower bound $\tau_m \gg \varepsilon_m^{1-\alpha-\gamma}$ (expected from (3.8)), and the upper bound $\tau_m \ll a_{m-1}^{-1}$ (implied by (3.12)), we choose $\tau_m^{-1} \gg 1$ to be an integer lying in the interval $(\varepsilon_{m-1}^{\alpha-1}, \varepsilon_m^{\alpha-1+\gamma})$.³² With τ_m chosen, we then return to (3.12) and let $(\tau_{m-1}'')^{-1}$ to be an integer lying in the interval (a_{m-1}, τ_m^{-1}) .³³ With these choices, the stream functions $\{\phi_m\}_{m \geq 0}$ are

²⁹If one wishes to prove Theorem 1.2, the parameters a_m appearing in (3.10) need to become “slow functions” of space and time, oscillating on scales ε_{m-1} in x and τ_{m-1} in t . The purpose of these slow functions in a convex-integration scheme is to cancel (through the mean of their squares) a leftover Euler-Reynolds stress error. We refer the reader to [16, Section 5] and [20, Section 7.3] for these details.

³⁰For technical reasons, the time-cutoff functions ζ and $\widehat{\zeta}$ are chosen to satisfy a number of fine properties not mentioned here, in order to deal with the fast switching of shear flows and the slow switching of flow maps; see [9, Section 2.1] for precise descriptions.

³¹More than that, this new time oscillation must itself be homogenized!

³²The choice $\tau_m^{-1} \in \mathbb{N}$ is made in order to ensure $[0, 1]$ -time periodicity of the respective time cutoffs. For instance, $\tau_m^{-1} \simeq \varepsilon_m^{\frac{(\alpha-1)(q+1)+\gamma q}{2q}}$ is an allowable choice.

³³For instance, an integer which satisfies $(\tau_m'')^{-1} \simeq \varepsilon_m^{\frac{(\alpha-1)(q+3)+\gamma q}{4}}$ is an allowable choice.

defined by the telescoping relation (3.10). The parameter $\Lambda = \varepsilon_1^{-1} \gg 1$ is chosen last, to absorb all m -independent constants appearing in our estimates.

The vector field u is obtained as $u = \lim_{m \rightarrow \infty} u_m$, the limit being taken with respect to the $C_{x,t}^{\alpha'}$ topology, for any $\alpha' < \alpha$. Here, $u_m = \nabla^\perp \phi_m$, $\phi_m = \sum_{1 \leq m' \leq m} (\phi_{m'} - \phi_{m'-1})$, $\phi_0 = 0$, and the stream function increment is given by (3.10).

Remark 3.1 (Genericity of the vector fields u from Theorem 1.2). The choice ϕ_0 is arbitrary, made in [9] only for notational simplicity. The proof of Theorem 1.2 works “as is” if we replace ϕ_0 by any sufficiently smooth, zero-mean periodic stream function. Then, all we have to do is to ensure that $\varepsilon_1 = \Lambda^{-1}$ is chosen to be (much) smaller than the smallest scale of oscillation of this chosen smooth ϕ_0 . By telescoping the available bounds on $\|\phi_m - \phi_{m-1}\|_{C_{x,t}^1}$ (see [9, Proposition 2.2]), we may then show that $\|\phi - \phi_0\|_{C_{x,t}^1} = \|u - \nabla^\perp \phi_0\|_{C_{x,t}^0} \leq C \Lambda^{-\alpha}$, where $C = C(\alpha, q) > 0$; this upper bound may be made arbitrarily small by choosing Λ sufficiently large, showing that the class of vector fields u for which Theorem 1.2 holds is in fact *dense*, in the class of all smooth divergence-free zero-mean vector fields, with respect to the $C_{x,t}^0$ topology. A stronger form of genericity of the family of vector fields $u \in C_{x,t}^\alpha$ which induce anomalous diffusion is obtained in [20], see Theorem 1.3 above.

One may show, see [9, Section 2.3] that the family $\{\phi_m\}_{m \geq 0}$ is uniformly bounded in $C_{x,t}^{1,\alpha}$, that for each m the stream function ϕ_m is real-analytic in space (uniformly in time), with radius of analyticity proportional to ε_m , that the flow maps X_m and their inverses X_m^{-1} remain close to the identity on their natural timescale τ_m'' , and that they are also real-analytic with radius of analyticity proportional to ε_m . That is to say, the family of stream functions $\{\phi_m\}_{m \geq 0}$ behave in the way we expect them to behave.

Remark 3.2 (Propagation of smoothness in the construction). In [9] we have chosen to prove the quantitative real-analyticity of the stream functions ϕ_m , of the Lagrangian flows X_m , and of their inverses X_m^{-1} , simply because it is true: the Ψ_k are sine waves, which are certainly analytic functions, the Lagrangian flow maps of an analytic function are analytic, adding and composing analytic functions to each other preserves analyticity; one has to be careful however when estimating the radius of analyticity of the resulting objects [9, Appendix B]. This real-analyticity is however not essential for the proof: all that is needed is “sufficiently high regularity”, to justify certain high-order expansions in the proof. Instead of working with real-analytic building blocks Ψ_k (sine waves), it is sufficient to consider smooth ones, and then to add a “mollification step” to the proof, as is common in convex-integration schemes (see e.g. [16, Section 2.4], [19, Section 5.3]). This perspective needs to be taken if one is to work with building blocks Ψ_k which are compactly supported in space, and this is precisely the perspective taken in the proof of Theorem 1.3, see [20, Section 7.1].

3.2. A backwards cascade of eddy diffusivities and homogenized problems. With the vector field u constructed in §3.1 it remains to rigorously justify our scale-by-scale homogenization picture, which dictates that when “coarse-grained” at scale ε_m , the advection-diffusion equation (1.1) with vector field u experiences an effective diffusivity κ_m , which satisfies the recursion relation (3.6), initialized with $\kappa_M = \kappa$ for a suitably chosen $M \geq 0$. This requires the introduction of a *permissible set of diffusivities* (the set \mathcal{K} mentioned in Theorem 1.2), and a definition of “coarse-graining” (at least, as it pertains to the analysis discussed here).

In spite of the inherent instability of the recursion formula $\kappa \mapsto \kappa + \frac{a_m^2 \varepsilon_m^4}{8\kappa}$, see (3.6), we have seen earlier in (3.7) that the relation $\kappa_m \simeq \varepsilon_m^{1+\alpha+\gamma}$ is consistent with stably solving (3.6). As such, we define the permissible set of diffusivities by

$$\mathcal{K} := \cup_{m \geq 0} \left[\frac{1}{2} \varepsilon_m^{1+\alpha+\gamma}, 2 \varepsilon_m^{1+\alpha+\gamma} \right].$$

For any given $\kappa \in \mathcal{K}$, we may identify a unique integer $M \geq 0$ such that

$$\kappa \in \left[\frac{1}{2} \varepsilon_M^{1+\alpha+\gamma}, 2 \varepsilon_M^{1+\alpha+\gamma} \right], \quad \text{and we define} \quad \kappa_M := \kappa. \quad (3.14)$$

Then, for $m \in \{1, \dots, M\}$, we recursively define the *renormalized diffusivities*

$$\kappa_{m-1} := \kappa_m \left(1 + c_* \left(\frac{a_m \varepsilon_m^2}{\kappa_m} \right)^2 \right). \quad (3.15)$$

Here, $c_* > 0$ is a universal constant, which may differ from the $\frac{1}{8}$ value presented in (3.6) depending on the L_t^2 -normalization of the time cutoffs ζ and $\widehat{\zeta}$.

Note that (3.14)–(3.15) defines a *backwards cascade of eddy diffusivities*,³⁴ which is $O(1)$ at the unit scale, in the sense that

$$\kappa = \kappa_M \ll \kappa_{M-1} \ll \dots \ll \kappa_m \ll \kappa_{m-1} \ll \dots \ll \kappa_0 \in \left[\frac{1}{2}, 2 \right].$$

³⁴It is verified in [9, Lemma 3.4] that $\kappa_m \in \mathcal{K}$ for all $m \in \{0, \dots, M\}$.

The above relation holds no matter how small the molecular diffusivity $0 < \kappa \ll 1$ is; except that as $\kappa \rightarrow 0$ we have $M \rightarrow \infty$, and thus in this limit we need to perform infinitely many renormalization/quantitative homogenization steps.

It is useful to keep in mind the relations

$$\frac{a_m \varepsilon_m^2}{\kappa_m} \simeq \varepsilon_m^{-\gamma} \gg 1, \quad \frac{\varepsilon_m^2}{\kappa_m \tau_m} \simeq \varepsilon_m^{\frac{(1-\alpha(1+2q))(q-1)}{2q(q+1)}} \ll 1,$$

which hold since $\kappa_m \simeq \varepsilon_m^{1+\alpha+\gamma} \simeq a_m \varepsilon_m^2 \cdot \varepsilon_m^\gamma$, since τ_m is given by the asymptotic in Footnote 32, and since $\alpha(2q+1) < 1$.

Having defined the permissible set of diffusivities \mathcal{K} , and the sequence $\{\kappa_m\}_{m=0}^M$ (see (3.14)–(3.15)), it thus remains to define what we mean by “coarse-graining” the advection-diffusion equation (1.1) at scale ε_m . A precise and robust definition of this concept, in a much more general setting is given in [4, 6]. For the purposes of proving Theorem 1.2, we are able to get away with a simpler perspective based on *iterated homogenization*, available since the construction of the field u relies solely on space-time periodic building blocks.

Fix $\kappa = \kappa_M \in \mathcal{K}$, and define θ^κ to be the solution the advection-diffusion equation (1.1) with drift velocity u and smooth initial condition θ_{in} . Then, for each $m \in \{0, \dots, M\}$, define θ_m to be the solution of the advection-diffusion equation with the same initial datum, but with *drift velocity given by u mollified at scale ε_m* (which in our context means that the velocity is taken to equal u_m), and with diffusivity parameter κ_m (as defined by (3.14)–(3.15)). That is, θ_m is defined as the solution of

$$\partial_t \theta_m + u_m \cdot \nabla \theta_m - \kappa_m \Delta \theta_m = 0, \quad \theta_m|_{t=0} = \theta_{\text{in}}, \quad (3.16)$$

for each $m \in \{0, \dots, M\}$. With this notation, we say that

the operator $\partial_t + u \cdot \nabla - \kappa \Delta$ “coarse-grains” at scale ε_m to the operator $\partial_t + u_m \cdot \nabla - \kappa_m \Delta$,

if we are able to show that we have

$$\frac{1}{C_*} \cdot \kappa_m \|\nabla \theta_m\|_{L_{t,x}^2}^2 \leq \kappa \|\nabla \theta^\kappa\|_{L_{t,x}^2}^2 \leq C_* \cdot \kappa_m \|\nabla \theta_m\|_{L_{t,x}^2}^2, \quad (3.17)$$

for some universal constant $C_* > 1$ (independent of m). If we are able to prove (3.17), then since $u_0 = 0$ and $\kappa_0 \simeq 1$, by the energy balance for (3.16) with $m = 0$, and appealing to the Poincaré inequality, we have $\kappa_0 \|\nabla \theta_0\|_{L_{t,x}^2}^2 \simeq \varrho \|\theta_{\text{in}}\|_{L^2}^2$ for some $\varrho \in (0, \frac{1}{2}]$; this would then prove Theorem 1.2. The proof of (3.17) is described next.

3.3. The heart of the matter: one step of the homogenization iteration. In order to establish (3.17) we take an iterative approach: using quantitative homogenization (akin to §2.2) for each $m \in \{1, \dots, M\}$ we show that the amount of energy dissipation experienced by θ_m is comparable to the amount of energy dissipation experienced by θ_{m-1} . In this “one step” of the homogenization process we make an error, and we aim to prove that these errors are summable in m , leading to (3.17).

Lemma 3.3 (below) summarizes this “one step” of homogenization induction, and this is the key step in the proof of Theorem 1.2. Its proof consists of comparing the equations for θ_m and θ_{m-1} by homogenizing the oscillations at scale ε_m . The fact that we can only get away with worrying about one scale at each step of the proof is because the scales ε_m are supergeometric, and thus becoming more well-separated as they get smaller: (3.1) shows that $\varepsilon_m/\varepsilon_{m-1} \simeq \varepsilon_{m-1}^{q-1} \rightarrow 0$ as $m \rightarrow \infty$. This is an undesirable and non-physical feature of our model, but is necessary for such a “naive” homogenization procedure to work, since we rely on the summability (in m) of the errors we make. In more realistic examples, one should expect to see true geometric scale separation, and this necessitates examining the interaction between *many* scales at each step of a renormalization iteration, and dealing with errors which are always of order one. There has been some recent progress [4, 6] in developing a general *quantitative theory of coarse-graining* for elliptic and parabolic operators, which is based on quantitative homogenization theory and is capable of analyzing such models—using a strategy which is broadly similar to our proof of Theorem 1.2.

The precise statement of the main homogenization estimate in our induction is given in [9, Proposition 5.2]; here we offer a simplified version (see also Remark 3.4):

Lemma 3.3 (The homogenization iteration). *Fix $\alpha \in (0, \frac{1}{3})$. There exists a small parameter $\delta = \delta(\alpha) \approx (q-1)^2 > 0$ and a universal constant $C = C(\alpha) > 0$, such that if $\varepsilon_1 = \Lambda^{-1}$ is sufficiently small (in terms of α alone) then the following statement holds. Assume the zero-mean initial datum θ_{in} is real-analytic, with analyticity radius $R_{\theta_{\text{in}}} \in (0, 1]$ which satisfies $R_{\theta_{\text{in}}} \geq \varepsilon_1^{1+\frac{\gamma}{2}}$ (recall, γ is defined in (3.7)). For $\kappa \in \mathcal{K}$, define M via (3.14). Then, for all $m \in \{2, \dots, M\}$ the solutions of (3.16) satisfy*

$$\|\theta_m - \theta_{m-1}\|_{L^\infty([0,1];L^2(\mathbb{T}^2))}^2 + \kappa_m \|\nabla \theta_m - \nabla \tilde{\theta}_m\|_{L^2([0,1] \times \mathbb{T}^2)}^2 \leq C \varepsilon_{m-1}^{2\delta} \|\theta_{\text{in}}\|_{L^2(\mathbb{T}^2)}^2 \quad (3.18)$$

and

$$\left| \frac{\kappa_m \|\nabla \theta_m\|_{L_{t,x}^2}^2}{\kappa_{m-1} \|\nabla \theta_{m-1}\|_{L_{t,x}^2}^2} - 1 \right| \leq C \varepsilon_{m-1}^\delta, \quad (3.19)$$

for a suitably defined function $\tilde{\theta}_m$ (see (3.20) below). Here $L_{t,x}^2 = L^2([0,1] \times \mathbb{T}^2)$.

Remark 3.4 (All real-analytic data). We show in [9, Proposition 5.2] that Lemma 3.3 holds for all zero-mean real-analytic data θ_{in} ,³⁵ no matter how small their analyticity radius³⁶ $R_{\theta_{\text{in}}}$ is. The only difference is that the bounds in (3.18) and (3.19) hold only for $m \in \{m_* + 1, \dots, M\}$, where $m_* := \min\{m \in \mathbb{N} : m \geq 1, R_{\theta_{\text{in}}} \geq \varepsilon_m^{1+\frac{\gamma}{2}}\}$.

Discussion of the proof of Lemma 3.3. The proof of this lemma is the bulk of the paper [9]. It requires solving separate space and a time homogenization problems, due to the three different “fast” scales in the equation for θ_m : the smallest spatial scale is ε_m , which is the length scale of oscillation of the shear flows; the smallest time scale, on which the shear flows switch directions, is given by τ_m ; lastly, we have the time scale τ_{m-1}'' on which the Lagrangian flow maps $X_{m-1,\ell}$ must refresh. We should think of these three fast scales as being well-separated, with the spatial scale ε_m being the smallest/fastest, and with $\tau_m \ll \tau_{m-1}''$ (see (3.11)).

The idea is the same as for the classical periodic case presented in §2.1 and §2.2. We want to make a two-scale ansatz $\tilde{\theta}_m$, similar to the one for $\tilde{\theta}_\varepsilon$ in (2.7). Here, the “macroscopic” function is θ_{m-1} , and akin to (2.8)–(2.9) we want to plug this two-scale ansatz into the “microscopic” equation, which is the equation for θ_m (see (3.16)). The hope is that the resulting error is small in $L_t^2 H_x^{-1}$, so that we can obtain good estimates by appealing to parabolic regularity theory, as in the proof of Lemma 2.1.

The initial “naive” idea would be to define $\tilde{\theta}_m$ by the two-scale ansatz

$$\tilde{\theta}_m = \theta_{m-1} + \sum_{k,\ell \in \mathbb{Z}} \xi \left(\frac{\cdot}{\tau_m} - k \right) \widehat{\xi} \left(\frac{\cdot}{\tau_{m-1}''} - \ell \right) \sum_{i \in \{1,2\}} (\chi_{m,k,e_i} \circ X_{m-1,\ell}^{-1}) \partial_i \theta_{m-1} + H_m,$$

where the χ_{m,k,e_i} are fast periodic “correctors” with slope e_i , which are oscillating at scale ε_m and are obtained by solving a cell-problem related to the shear flow Ψ_k ; the H_m is responsible for homogenizing the time cutoffs, and the $\xi, \widehat{\xi}$ are unit-scale time cutoffs that are slightly wider than ζ and $\widehat{\zeta}$. Unfortunately, compared to the periodic setting, we face a number of additional complications and for these reasons, the definition above is not precise enough and cannot work. First, our equation is not ε_m -periodic because of “macroscopic” perturbations from the scales $\varepsilon_{m-1}, \dots, \varepsilon_0$, including those caused by the Lagrangian flows. Moreover, we cannot just “freeze” the coefficients and make an error here; some of these macroscopic perturbations need to be homogenized along with the ε_m -scale oscillations. For these reasons, our two-scale ansatz is a bit more complicated.

The *actual* definition of $\tilde{\theta}_m$ which is used in (3.18) is given in [9, (4.24)] as

$$\tilde{\theta}_m = T_{m-1} + \sum_{k,\ell \in \mathbb{Z}} \xi \left(\frac{\cdot}{\tau_m} - k \right) \widehat{\xi} \left(\frac{\cdot}{\tau_{m-1}''} - \ell \right) \sum_{i \in \{1,2\}} \tilde{\chi}_{m,k,e_i} (\nabla (T_{m-1} \circ X_{m-1,\ell}) \circ X_{m-1,\ell}^{-1})_i + \tilde{H}_m, \quad (3.20)$$

where:

- The “twisted correctors” $\tilde{\chi}_{m,k,e_i}$ play the role of the fast periodic correctors with slope e_i . These are simply the periodic correctors for the horizontal and vertical shears at scale ε_m , denoted by χ_{m,k,e_i} (see [9, (3.6)]), evaluated along the inverse flow $X_{m-1,k}^{-1}$ (see [9, (4.26)]). Note that in (3.20) the twisted correctors $\tilde{\chi}_{m,k,e_i}$, which are indexed by $k \in \mathbb{Z}$, are simply glued together in time by the partition of unity given by $\xi(\frac{\cdot}{\tau_m} - k)$; all other terms in (3.20) are indexed by $\ell \in \mathbb{Z}$.
- The \tilde{H}_m is an approximate time corrector, see the below discussion and [9, (4.21)].
- The function T_{m-1} (defined in [9, (4.13)]) is a perturbation of the function θ_{m-1} . It is chosen to be an *approximate* solution of the equation

$$\partial_t T_{m-1} + u_{m-1} \cdot \nabla T_{m-1} - \nabla \cdot (\mathbf{K}_m + \mathbf{s}_{m-1}) \nabla T_{m-1} = 0, \quad T_{m-1}|_{t=0} = \theta_{\text{in}}, \quad (3.21)$$

where

- The field $\mathbf{K}_m(t)$ (see [9, (3.26)]) is a time-oscillating field that oscillates between the homogenized matrices for the vertical and horizontal shears; it averages to $\kappa_{m-1} \mathbf{I}_2$, plus a negligible error.

³⁵Note that Theorem 1.2 is stated for $H^1(\mathbb{T}^d)$ -smooth initial data θ_{in} , not real-analytic one. However, upon mollification any $H^1(\mathbb{T}^d)$ data can be well-approximated by a real-analytic one, and this suffices for the proof to close; see §3.4.

³⁶We say that $f: \mathbb{T}^d \rightarrow \mathbb{R}$ is real-analytic with analyticity radius (at least) $R_f \in (0, 1]$ if $\sup_{n \geq 0} \|\nabla^n f\|_{L^2(\mathbb{T}^d)} R_f^n (n!)^{-1} \leq \|f\|_{L^2(\mathbb{T}^d)}$.

- The field s_{m-1} is defined as

$$s_{m-1} := \mathbf{K}_m \sum_{\ell \in \mathbb{Z}} \widehat{\xi} \left(\frac{\cdot}{\tau_{m-1}''} - \ell \right) (\nabla X_{m-1,\ell} \circ X_{m-1,\ell}^{-1} - \mathbf{I}_2).$$

It is added to the diffusion matrix of the T_{m-1} equation in order to deal with “distortions” caused by the Lagrangian flows; these distortions are small, but cannot be neglected in the equation for θ_m .

Equation (3.21) is close to the equation for θ_{m-1} (see (3.16) with m replaced by $m-1$) in the sense that s_{m-1} is small (due to (3.9)) and the time oscillating field $\mathbf{K}_m(t)$ will homogenize to $\kappa_{m-1}\mathbf{I}_2$, since time oscillations simply average.

A warning: for technical reasons, we do *not* define the function T_{m-1} to be a solution of equation (3.21)! The reason is that equation (3.21) does not possess sufficiently strong regularity estimates; this is because when compared to $\kappa_{m-1}\mathbf{I}_2$, the matrix \mathbf{K}_m has ellipticity constants which are much worse. Our regularity estimates for T_{m-1} must be very sharp, since it is the “macroscopic” function here and, as we saw in §2.1 and §2.2, higher derivatives of the macroscopic function show up in the quantitative homogenization estimates. To fix this problem, we choose T_{m-1} to be an approximate solution of (3.21) which we construct by hand, using a high-order iterative scheme. We cut off the iteration defining T_{m-1} at a finite stage, ensuring that the regularity estimates we need to hold are valid, while making a very small error in the equation for T_{m-1} . It is in these high-order estimates for T_{m-1} that we use the quantitative real-analyticity assumption of the initial data present in the statement of Lemma 3.3 (we refer the reader to [9, Section 4.3]).

Similarly, the function \tilde{H}_m in (3.20) plays the role of the time corrector (analogous to h_e in §2.1). However, as for T_{m-1} , in order to ensure that we have good enough regularity estimates for \tilde{H}_m , we actually construct it as an approximate time corrector using an iterative scheme, to ensure that it has better regularity properties.

With the definition of the ansatz (3.20) in hand, the argument in [9] proceeds by plugging $\tilde{\theta}_m$ into the equation for θ_m , and checking that the right-hand side is small. The error is quite explicit, although we do not write it here since it takes half a page to display the nine error terms: see [9, (5.18)–(5.26)]. It is then a tedious task to check that all these terms are suitably small in $L_t^2 H_x^{-1}$, but after doing so we obtain the bounds (3.18)–(3.19) claimed in the lemma. \square

3.4. The proof of Theorem 1.2: anomalous diffusion. To conclude this section, we show how Lemma 3.3 (and Remark 3.4) imply that the family of solutions $\{\theta^\kappa\}_{\kappa \in \mathcal{K}}$ of the advection-diffusion equation (1.1) with velocity u , exhibits anomalous diffusion in the sense of Definition 1.1, as promised in the statement of Theorem 1.2. For the proofs of the statement made in bullets three–six in Theorem 1.2, we refer the reader to [9].

Assume for the moment that the function θ_{in} is real-analytic.³⁷ Denote by $R_{\theta_{\text{in}}} \in (0, 1]$ its analyticity radius, so that $\sup_{n \geq 0} \|\nabla^n \theta_{\text{in}}\|_{L^2(\mathbb{T}^d)} R_{\theta_{\text{in}}}^n (n!)^{-1} \leq \|\theta_{\text{in}}\|_{L^2(\mathbb{T}^d)}$. Define $m_* \geq 1$ such that $\varepsilon_{m_*}^{1+\frac{\gamma}{2}} \simeq R_{\theta_{\text{in}}}$; the precise definition of m_* is given in Remark 3.4. This definition, together with the relation $\kappa_m \simeq \varepsilon_m^{1+\alpha+\gamma}$ for all $m \geq 0$ (see (3.7)), gives

$$\kappa_{m_*} \simeq \varepsilon_{m_*}^{1+\alpha+\gamma} \simeq R_{\theta_{\text{in}}}^{\frac{2(1+\alpha+\gamma)}{2+\gamma}}. \quad (3.22)$$

Relation (3.22) shows that κ_{m_*} is a function of $R_{\theta_{\text{in}}}$ alone (it is independent of the molecular diffusivity parameter κ). By the energy inequality for the θ_{m_*} evolution (that is, (3.16) with m replaced by m_*), and the Poincaré inequality, we deduce

$$\frac{1}{2} \frac{d}{dt} \|\theta_{m_*}(t, \cdot)\|_{L^2(\mathbb{T}^2)}^2 = -\kappa_{m_*} \|\nabla \theta_{m_*}(t, \cdot)\|_{L^2(\mathbb{T}^2)}^2 \leq -(2\pi)^2 \kappa_{m_*} \|\theta_{m_*}(t, \cdot)\|_{L^2(\mathbb{T}^2)}^2.$$

In turn, the above estimate implies

$$\kappa_{m_*} \|\nabla \theta_{m_*}\|_{L^2([0,1] \times \mathbb{T}^2)}^2 = \frac{1}{2} \|\theta_{\text{in}}\|_{L^2(\mathbb{T}^2)}^2 - \frac{1}{2} \|\theta_{m_*}(1, \cdot)\|_{L^2(\mathbb{T}^2)}^2 \geq \frac{1}{2} \|\theta_{\text{in}}\|_{L^2(\mathbb{T}^2)}^2 (1 - e^{-8\pi^2 \kappa_{m_*}}) \geq 2\kappa_{m_*} \cdot \frac{1}{2} \|\theta_{\text{in}}\|_{L^2(\mathbb{T}^2)}^2. \quad (3.23)$$

Here we have used that without loss of generality $\kappa_{m_*} \leq \frac{1}{3}$ (since $m_* \geq 1$, and we can take Λ to be sufficiently large), and so $1 - e^{-8\pi^2 \kappa_{m_*}} \geq 2\kappa_{m_*}$.

Next, take any $\kappa \in \mathcal{K}$, and define $M = M(\kappa) \geq 0$ according to (3.14), so that $\kappa = \kappa_M \simeq \varepsilon_M^{1+\alpha+\gamma}$; in particular, as $\kappa \rightarrow 0$ we have $M \rightarrow \infty$, and so $M > m_*$ for all sufficiently small κ (since m_* is only a function of $R_{\theta_{\text{in}}}$, which is

³⁷At the end of the proof we will use the fact that $H^1(\mathbb{T}^2)$ functions can be approximated by real-analytic ones, with an error that can be made quantitatively small in $L^2(\mathbb{T}^2)$.

fixed). By Remark 3.4, the bounds (3.18)–(3.19) hold for all m in the nontrivial range $\{m_* + 1, \dots, M\}$. For instance, by telescoping estimate (3.19), we obtain that for all $m \in \{m_*, \dots, M\}$

$$\prod_{m=m_*+1}^M (1 - C\varepsilon_{m-1}^\delta) \leq \frac{\kappa_m \|\nabla \theta_m\|_{L^2([0,1] \times \mathbb{T}^2)}^2}{\kappa_{m_*} \|\nabla \theta_{m_*}\|_{L^2([0,1] \times \mathbb{T}^2)}^2} \leq \prod_{m=m_*+1}^M (1 + C\varepsilon_{m-1}^\delta).$$

The super-geometric growth $\varepsilon_m \simeq \varepsilon_{m-1}^q$ with $q > 1$, together with the fact that $\varepsilon_1 = \Lambda^{-1}$ is chosen sufficiently small (in terms of α), implies

$$\frac{3}{4} \leq \prod_{m=2}^{\infty} (1 \pm C\varepsilon_{m-1}^\delta) \leq \frac{4}{3},$$

and thus for all $m \in \{m_*, \dots, M\}$, we have

$$\frac{3}{4} \kappa_{m_*} \|\nabla \theta_{m_*}\|_{L^2([0,1] \times \mathbb{T}^2)}^2 \leq \kappa_M \|\nabla \theta_M\|_{L^2([0,1] \times \mathbb{T}^2)}^2 \leq \frac{4}{3} \|\nabla \theta_{m_*}\|_{L^2([0,1] \times \mathbb{T}^2)}^2. \quad (3.24)$$

Estimate (3.24) is used with $M = m_*$ in the proof.

In view of the previously established lower bound (3.23), it remains to relate the term $\kappa_M \|\nabla \theta_M\|_{L^2([0,1] \times \mathbb{T}^2)}^2 = \kappa \|\nabla \theta_{M(\kappa)}\|_{L^2([0,1] \times \mathbb{T}^2)}^2$ appearing in (3.24), to the true energy dissipation rate appearing in (1.5), namely $\kappa \|\nabla \theta^\kappa\|_{L^2([0,1] \times \mathbb{T}^2)}^2$. In order to achieve this, we note that $\theta^\kappa - \theta_M$ satisfies the equation

$$(\partial_t - \kappa \Delta + u \cdot \nabla)(\theta^\kappa - \theta_M) = \operatorname{div}((\phi - \phi_M) \nabla \theta_M), \quad (\theta^\kappa - \theta_M)|_{t=0} = 0.$$

Since by construction we have $\|\phi - \phi_M\|_{L_{x,t}^\infty} \leq C\varepsilon_M^{(1+\alpha)q} \ll 1$ (which follows by telescoping the trivial bound for (3.10)), a simple energy estimate gives

$$\|\theta^\kappa - \theta_M\|_{L^\infty([0,1]; L^2(\mathbb{T}^2))}^2 + \kappa \|\nabla \theta^\kappa - \nabla \theta_M\|_{L^2([0,1] \times \mathbb{T}^2)}^2 \leq \frac{1}{8} \kappa \|\nabla \theta_M\|_{L^2([0,1] \times \mathbb{T}^2)}^2. \quad (3.25)$$

Combining (3.22), (3.23), (3.24), (3.25), and recalling that $\kappa = \kappa_M$, we deduce

$$\begin{aligned} \kappa \|\nabla \theta^\kappa\|_{L^2([0,1] \times \mathbb{T}^2)}^2 &\geq \kappa \|\nabla \theta_M\|_{L^2([0,1] \times \mathbb{T}^2)}^2 - \kappa \|\nabla \theta^\kappa - \nabla \theta_M\|_{L^2([0,1] \times \mathbb{T}^2)}^2 \\ &\geq \frac{7}{8} \kappa_M \|\nabla \theta_M\|_{L^2([0,1] \times \mathbb{T}^2)}^2 \geq \frac{7}{8} \cdot \frac{3}{4} \cdot \kappa_{m_*} \|\nabla \theta_{m_*}\|_{L^2([0,1] \times \mathbb{T}^2)}^2 \\ &\geq \kappa_{m_*} \cdot \frac{1}{2} \|\theta_{\text{in}}\|_{L^2(\mathbb{T}^2)}^2 \geq C_* R_{\theta_{\text{in}}}^{\frac{2(1+\alpha+\gamma)}{2+\gamma}} \cdot \frac{1}{2} \|\theta_{\text{in}}\|_{L^2(\mathbb{T}^2)}^2. \end{aligned} \quad (3.26)$$

In order to complete the proof of anomalous diffusion for all zero-mean data in $H^1(\mathbb{T}^2)$, we need to address the real-analyticity assumption used in (3.26), which explicitly contains the analyticity radius $R_{\theta_{\text{in}}}$.

For this purpose, for any $\theta_{\text{in}} \in H^1(\mathbb{T}^2)$ we define $L_{\text{in}} := \frac{\|\theta_{\text{in}}\|_{L^2}}{\|\nabla \theta_{\text{in}}\|_{L^2}}$. Then, we mollify θ_{in} at spatial scale λL_{in} , for a parameter λ which is to be determined; that is, we define

$$\tilde{\theta}_{\text{in}} := \theta_{\text{in}} * \Phi(\lambda^2 L_{\text{in}}^2, \cdot),$$

where $\Phi(t, x)$ is the standard heat kernel (with unit diffusion coefficient). The function $\tilde{\theta}_{\text{in}}$ is real-analytic, and its analyticity radius is given by $R_{\tilde{\theta}_{\text{in}}} = c\lambda L_{\text{in}}$, for a universal constant $c > 0$. Moreover, standard properties of the heat kernel Φ imply that

$$\|\theta_{\text{in}}\|_{L^2(\mathbb{T}^2)}^2 \geq \|\tilde{\theta}_{\text{in}}\|_{L^2(\mathbb{T}^2)}^2 \geq \|\theta_{\text{in}}\|_{L^2(\mathbb{T}^2)}^2 - C\lambda^2 L_{\text{in}}^2 \|\nabla \theta_{\text{in}}\|_{L^2(\mathbb{T}^2)}^2 \geq (1 - C\lambda^2) \|\theta_{\text{in}}\|_{L^2(\mathbb{T}^2)}^2. \quad (3.27)$$

To conclude, we denote by $\tilde{\theta}^\kappa$ the solution of the advection-diffusion equation (1.1) with initial condition $\tilde{\theta}_{\text{in}}$. By linearity, we have that $(\partial_t + u \cdot \nabla - \kappa \Delta)(\tilde{\theta}^\kappa - \theta^\kappa) = 0$, and so by the standard energy balance for this difference (recall (1.2)), the triangle inequality, the bound (3.26) with θ^κ replaced by $\tilde{\theta}^\kappa$ and θ_{in} replaced by $\tilde{\theta}_{\text{in}}$, and with the bounds in (3.27), we deduce

$$\begin{aligned} \kappa \|\nabla \theta^\kappa\|_{L^2([0,1] \times \mathbb{T}^2)}^2 &\geq \kappa \|\nabla \tilde{\theta}^\kappa\|_{L^2([0,1] \times \mathbb{T}^2)}^2 - \kappa \|\nabla(\tilde{\theta}^\kappa - \theta^\kappa)\|_{L^2([0,1] \times \mathbb{T}^2)}^2 \\ &\geq C_* R_{\tilde{\theta}_{\text{in}}}^{\frac{2(1+\alpha+\gamma)}{2+\gamma}} \cdot \frac{1}{2} \|\tilde{\theta}_{\text{in}}\|_{L^2(\mathbb{T}^2)}^2 - \frac{1}{2} \|\tilde{\theta}_{\text{in}} - \theta_{\text{in}}\|_{L^2(\mathbb{T}^2)}^2 \\ &\geq \frac{C_*}{2} (c\lambda L_{\text{in}})^{\frac{2(1+\alpha+\gamma)}{2+\gamma}} \cdot \frac{1}{2} \|\theta_{\text{in}}\|_{L^2(\mathbb{T}^2)}^2 - C\lambda^2 \cdot \frac{1}{2} \|\theta_{\text{in}}\|_{L^2(\mathbb{T}^2)}^2, \end{aligned} \quad (3.28)$$

assuming that $C\lambda^2 \leq \frac{1}{2}$. The bound (3.28) dictates that we choose λ small enough ensure that the first term on the right side dominates; this amounts to letting

$$\lambda \leq c_* L_{\text{in}}^{\frac{1+\alpha+\gamma}{1-\alpha}},$$

for some universal constant $0 < c_* \ll 1$ which depends on c , C , and C_* . Inserting the above bound into (3.28) implies

$$\kappa \|\nabla \theta^\kappa\|_{L^2([0,1] \times \mathbb{T}^2)}^2 \geq \underbrace{\left(\frac{C_*}{4} (cc_*)^{\frac{2(1+\alpha+\gamma)}{2+\gamma}} \right) \cdot L_{\text{in}}^{\frac{2(1+\alpha+\gamma)}{1-\alpha}} \cdot \frac{1}{2} \|\theta_{\text{in}}\|_{L^2(\mathbb{T}^2)}^2}_{=: \varrho}.$$

The dependence of the coefficient ϱ on L_{in} claimed in the first bullet of Theorem 1.2 now follows by letting $\varepsilon = \frac{2\gamma}{1-\alpha} = \frac{2(q-1)(1+\alpha)}{(q+1)(1-\alpha)} \rightarrow 0$ as $q \rightarrow 1$.

4. EXTENSIONS AND OPEN PROBLEMS

4.1. Uniform Hölder regularity of the scalars. An important question which is not addressed by Theorem 1.2 is the *regularity* of the solutions θ^κ , in particular whether there is a uniform-in- κ Hölder estimate such as

$$\sup_{\kappa \in \mathcal{K}} \|\theta^\kappa\|_{L^2([0,1]; C^{0,\bar{\alpha}}(\mathbb{T}^d))} < \infty. \quad (4.1)$$

If we could show (4.1) for $\bar{\alpha}$ arbitrarily close to $\frac{1-\alpha}{2}$, then this would resolve part (ii) of the Obukhov-Corrsin dichotomy discussed in §1.2.

To see why we should expect such an estimate, let us recall the Campanato characterization of the Hölder space $C^{0,\beta}$, which states that³⁸ for any $f \in H^1(\mathbb{T}^d)$,

$$[f]_{C^{0,\beta}(\mathbb{T}^d)} \simeq \sup_{r \in (0,1)} r^{1-\beta} \|\nabla(f * \eta_r)\|_{L^\infty(\mathbb{T}^d)}. \quad (4.2)$$

What we have shown in Theorem 1.2 (recall (3.24) and also (3.7)) is that

$$\|\nabla \theta_m\|_{L^2([0,1] \times \mathbb{T}^d)} \approx \kappa_m^{-\frac{1}{2}} \simeq \varepsilon_m^{-\frac{1+\alpha+\gamma}{2}},$$

where we ignore the dependence on the initial datum θ_{in} of the implicit constants in the \approx symbol, and we recall that $\gamma > 0$ is a very small exponent (proportional to $q-1$, which can thus be made arbitrarily small in our construction of the vector field). While not explicitly stated, the estimate in Lemma 3.3 implies that, for $j < m$, we have

$$\kappa_j \|\nabla \theta_m * \eta_{\varepsilon'_j, \tau''_j} - \nabla \theta_j\|_{L^2([0,1] \times \mathbb{T}^d)}^2 \lesssim \varepsilon_{j-1}^{2\delta}, \quad (4.3)$$

where $\varepsilon'_j = \sqrt{\varepsilon_{j+1} \varepsilon_j}$ and $\tau''_j := \sqrt{\tau_j \tau'_j}$ are chosen so that $\varepsilon_j < \varepsilon'_{j'} < \varepsilon_{j-1}$ and $\tau'_j < \tau''_j < \tau_j$, with some amount of separation between each of these scales; again, we ignore the dependence on the initial datum θ_{in} of the implicit constants in the \lesssim symbol. In other words, we should consider $\nabla \theta_j \simeq \nabla \theta_m * \eta_{\varepsilon'_j, \tau''_j}$ if $j < m$. Therefore the previous two displays imply the bound

$$\|\nabla(\theta_m * \eta_{\varepsilon'_j, \tau''_j})\|_{L^2([0,1] \times \mathbb{T}^d)} \approx \kappa_j^{-\frac{1}{2}} \simeq \varepsilon_j^{-\frac{1+\alpha+\gamma}{2}} \simeq (\varepsilon'_j)^{-\frac{1+\alpha}{2} - O(\gamma)}$$

If we could upgrade this bound from L^2 to L^∞ ,³⁹ then by (4.2) we would have the desired uniform Hölder estimate for θ_m with regularity exponent $\bar{\alpha} = 1 - \frac{1+\alpha}{2} - O(\gamma) = \frac{1-\alpha}{2} - O(\gamma)$.

The aforementioned “upgrade” amounts to proving the Lipschitz-type estimate

$$\|\nabla \theta_m\|_{L_t^2 L_x^\infty([0,1] \times \mathbb{T}^d)} \approx \kappa_m^{-\frac{1}{2}}, \quad (4.4)$$

which states that the gradient field $\nabla \theta_m$ does not concentrate on sets of small measure: it is roughly the same size, everywhere in \mathbb{T}^d . It is very natural to expect such an estimate to be true. First, at an intuitive level, since: (i) our vector field is built from periodic ingredients with supergeometric scale separation between successive scales (so that the main contribution of $\nabla \theta_m$ are the wiggles at scale ε_m), and (ii) we have shown that the solutions have expansions with periodic ingredients, there is no reason to expect any such “concentration” to occur.

Second, uniform Lipschitz estimates of exactly this type have played a central role in homogenization theory since the pioneering work of Avellaneda and Lin [10, 11] in the ’80s; they now go by the name *large-scale regularity theory* (see the monographs [7, 65] and the references therein). While the arguments of Avellaneda and Lin [10, 11] are based on compactness and apply only to equations with periodic coefficients and thus one active scale, a quantitative approach to large-scale regularity was later proposed in [8]; this method is more robust and applies to equations with non-periodic coefficients (such as almost periodic and random coefficients).

³⁸We denote by $\{\eta_r\}_{r>0}$ a standard family of mollifiers, with η_r localized to scale $\simeq r$, and by $\{\eta_{r,s}\}_{r,s>0}$ a family space-time mollifiers at spatial scale r and time scale s .

³⁹We would also need some uniform-in-time estimates to treat the time mollification.

The way it works heuristically is that the regularity of the homogenized equation is “transferred” to the equation with oscillating coefficients using quantitative homogenization estimates, and an excess decay iteration. In our setting, the “homogenized equation” would be the equation for θ_{m-1} and the “equation with oscillating coefficients” would be the equation for θ_m . This sets us an induction *going down the scales* in which the regularity for θ_{m-1} is transferred to θ_m : a cascade of regularity, dual to the inverse cascade of homogenization. Such an iteration argument would obviously be very technical to implement and take a great deal of effort to write, but we have all the ingredients to implement it in our setting,⁴⁰ and we do not expect that formalizing it would require major new ideas beyond [8, 9].

As part of the induction argument that proves (4.4), we will also upgrade the estimate (4.3) to an L^∞ -type estimate, and thus the uniform-in- m Hölder estimate for θ_m can be obtained.

One nice consequence of such a Hölder estimate would be that the condition in Theorem 1.2 that the initial data belong to $H^1(\mathbb{T}^d)$ would be removed, and the positive constant ϱ would be universal. This would then imply the exponential decay in time of the $L^\infty(\mathbb{T}^d)$ -norm of our solutions θ^κ , uniformly along our subsequence κ_j .

4.2. Geometric separation of scales and intermittency. A major shortcoming of the vector field constructed in Section 3.1 is that the active scales $\{\varepsilon_m\}_{m \geq 0}$ are supergeometrically separated. This nonphysical aspect was necessary for our proof of anomalous diffusion, based on iterative quantitative homogenization, to work. If we want the methods in [9] to apply to more realistic physical models, we need to develop a more flexible method.

Therefore, we think that it is an important open problem is to prove anomalous diffusion for a variant of our construction in which *the scales $\{\varepsilon_m\}_{m \geq 0}$ are geometrically separated*, i.e., $\varepsilon_m = \exp(-Cm)$.

If the scales are geometrically separated, then we would make *an error of order one at each scale m* —which we obviously cannot sum up over m ! Moreover, we would have to deal with leading-order interactions between multiple scales at once. In contrast, the proof of Theorem 1.2 is based on comparing the equation for θ_{m-1} to the equation for θ_m ; evidently the scales ε_{m+1} and ε_{m-1} do not interact directly, which would not be the case if they were geometrically separated.

It is only due to the supergeometric scale separation that we should expect the Hölder estimate discussed in Section 4.1 above to be valid. A geometric separation of scales—which implies the interaction between many scales at once—should lead to concentration effects and therefore intermittency. We would expect in this case that the L^p norm of $\nabla\theta_m$ to be very different for different values of p .

At first glance, these issues may seem fatal to our whole strategy based on homogenization theory. However, a general “coarse-graining” theory, based on quantitative homogenization methods, has been recently developed [5, 6]. Based on this, and on an iterative quantitative homogenization argument, a superdiffusive central limit theorem was recently proved in [4] for a passive scalar model with a random vector field (with non-separated scales). There is therefore some hope that iterative quantitative homogenization methods as in [4, 9] may apply to more physically realistic models of scalar turbulence.

4.3. A regularity threshold for Euler solutions exhibiting anomalous diffusion? We conjecture that the regularity threshold $\alpha = \frac{1}{3}$ present in Theorem 1.2 is not an artifact of our proof in [9], but rather, a delicate rigidity constraint imposed by the Lagrangian nature of turbulent diffusion. We propose the following problem:

Conjecture 4.1 (Anomalous diffusion + self advection = Onsager super-criticality). *Let $d \in \{2, 3\}$. Assume that $u \in C^0([0, 1]; C^\alpha(\mathbb{T}^d)) \cap C^\alpha([0, 1]; C^0(\mathbb{T}^d))$ is a weak solution of the incompressible Euler equations, for some Hölder regularity exponent $\alpha \in (0, 1)$. Furthermore, assume that for any initial condition $\theta_{\text{in}} \in \dot{H}^1(\mathbb{T}^d)$, the family of solutions $\{\theta^\kappa\}_{\kappa > 0}$ of the passive scalar equation (1.1) with velocity field u displays anomalous diffusion continuously in time; i.e., (1.5) holds and the time dissipation measure $\mathcal{E}(dt)$ from (1.6) is non-atomic. Then, we have that $\alpha \leq \frac{1}{3}$.*

We expect that the method of proof for resolving Conjecture 4.1 is more interesting than what the statement of the Conjecture yields. Indeed, if one wishes to prove that the Conjecture is *false*, then one first needs to construct $C_{x,t}^\alpha$ weak solutions of the incompressible Euler equations, for some $\alpha \in (\frac{1}{3}, 1)$. This question is widely open (in the absence of artificial forcing terms). Conversely, if one is to prove that the Conjecture is *true*, then one needs to synthesize weak Euler solutions from the information gained by watching what the associated advection-diffusion equation does to arbitrary $H^1(\mathbb{T}^d)$ smooth scalar initial conditions, in the vanishing diffusivity limit; one of the clear enemies is the pressure term in the Euler system.

⁴⁰We would need a localized version of Lemma 3.3, which applies not just in the whole torus \mathbb{T}^2 , but in all appropriately scaled space-time cylinders above the scale at which we expect homogenization to occur. Such an estimate can be obtained by an argument similar to that of Lemma 3.3, there will just be additional boundary layer terms to control.

ACKNOWLEDGEMENTS

S.A. was supported by NSF grant DMS-2350340. V.V. was in part supported by the Collaborative NSF grant DMS-2307681 and a Simons Investigator Award.

REFERENCES

- [1] G. Alberti, G. Crippa, and A. L. Mazzucato. Exponential self-similar mixing by incompressible flows. *J. Amer. Math. Soc.* **32** (2019), no. 2, 445–490.
- [2] L. Ambrosio. Transport equation and Cauchy problem for BV vector fields. *Invent. Math.* **158** (2004), no. 2, 227–260.
- [3] 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** (1956), no. 1200, 67–77.
- [4] S. Armstrong, A. Bou-Rabee, and T. Kuusi. Superdiffusive central limit theorem for a Brownian particle in a critically-correlated incompressible random drift. 2024, ArXIV preprint 2404.01115.
- [5] S. Armstrong, T. Kuusi. Elliptic homogenization from qualitative to quantitative. 2022, ArXIV preprint 2210.06488.
- [6] S. Armstrong, T. Kuusi. Renormalization group and elliptic homogenization in high contrast. 2024, ArXIV preprint 2405.10732.
- [7] S. Armstrong, T. Kuusi and J.-C. Mourrat. *Quantitative stochastic homogenization and large-scale regularity*. Grundlehren der mathematischen Wissenschaften, 352, Springer, Cham, 2019.
- [8] S. N. Armstrong and C. K. Smart. Quantitative stochastic homogenization of convex integral functionals. *Ann. Sci. Éc. Norm. Supér. (4)* **49** (2016), no. 2, 423–481.
- [9] S. Armstrong and V. Vicol. Anomalous Diffusion by Fractal Homogenization. *Annals of PDE* **11** (2025), no. 1, paper no. 2.
- [10] M. Avellaneda and F. Lin. Compactness methods in the theory of homogenization. *Comm. Pure Appl. Math.* **40** (1987), no. 6, 803–847.
- [11] M. Avellaneda and F. Lin. L^p bounds on singular integrals in homogenization. *Comm. Pure Appl. Math.* **44** (1991), no. 8-9, 897–910.
- [12] 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** (1992), no. 1, 139–204.
- [13] G.K. Batchelor. Small-scale variation of convected quantities like temperature in turbulent fluid Part 1. General discussion and the case of small conductivity. *Journal of fluid mechanics* **5** (1959), no. 1, 113–133.
- [14] D. Bernard, K. Gawedzki, and A. Kupiainen. Slow modes in passive advection. *Journal of Statistical Physics* **90** (1998), no. 3, 519–569.
- [15] E. Brûè and C. De Lellis. Anomalous dissipation for the forced 3D Navier-Stokes equations. *Comm. Math. Phys.* **400** (2023), no. 3, 1507–1533.
- [16] T. Buckmaster, C. De Lellis, L. Székelyhidi Jr., and V. Vicol. Onsager’s conjecture for admissible weak solutions. *Comm. Pure Appl. Math.* **72** (2019), no. 2, 229–274.
- [17] T. Buckmaster and V. Vicol. Nonuniqueness of weak solutions to the Navier-Stokes equation. *Ann. of Math. (2)* **189** (2019), no. 1, 101–144.
- [18] T. Buckmaster, M. Colombo, and V. Vicol. Wild solutions of the Navier-Stokes equations whose singular sets in time have Hausdorff dimension strictly less than 1. *J. Eur. Math. Soc. (JEMS)* **24** (2022), no. 9, 3333–3378.
- [19] T. Buckmaster and V. Vicol. Convex integration and phenomenologies in turbulence. *EMS Surveys in Mathematical Sciences* **6** (2020), no. 1, 173–263.
- [20] J. Burczak, L. Székelyhidi Jr., and B. Wu. Anomalous dissipation and Euler flows. 2024, ArXIV preprint 2310.02934.
- [21] A. Cheskidov. Dissipation anomaly and anomalous dissipation in incompressible fluid flows. 2023, ArXIV preprint 2311.04182.
- [22] A. Cheskidov, P. Constantin, S. Friedlander, and R. Shvydkoy. Energy conservation and Onsager’s conjecture for the Euler equations. *Nonlinearity* **21** (2008), no. 6, 1233–1252.
- [23] A. Cheskidov and X. Luo. Nonuniqueness of weak solutions for the transport equation at critical space regularity. *Annals of PDE* **7** (2021), no. 1, paper no. 2.
- [24] M. Colombo, G. Crippa, and M. Sorella. Anomalous dissipation and lack of selection in the Obukhov-Corrsin theory of scalar turbulence. *Annals of PDE* **9** (2023), no. 2, paper no. 21.
- [25] 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** (1994), no. 1, 207–209.
- [26] S. Corrsin. On the spectrum of isotropic temperature fluctuations in an isotropic turbulence. *J. Appl. Phys.* **22** (1951), 469–473.
- [27] S. Daneri and L. Székelyhidi, Jr. Non-uniqueness and h -principle for Hölder-continuous weak solutions of the Euler equations. *Arch. Ration. Mech. Anal.* **224** (2017), no. 2, 471–514.
- [28] C. De Lellis and L. Székelyhidi, Jr. The h -principle and the equations of fluid dynamics. *Bull. Amer. Math. Soc.* **49** (2012), no. 3, 347–375.
- [29] C. De Lellis and L. Székelyhidi Jr. Dissipative continuous Euler flows. *Invent. Math.* **193** (2013), no. 2, 377–407.
- [30] R.J. DiPerna and P.-L. Lions. Ordinary differential equations, transport theory and Sobolev spaces. *Invent. Math.* **98** (1989), no. 3, 511–547.
- [31] D.A. Donzis, K.R. Sreenivasan, and P.K. Yeung. Scalar dissipation rate and dissipative anomaly in isotropic turbulence. *Journal of Fluid Mechanics* **532** (2005), 199–216.
- [32] 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** (2017), 153–189.
- [33] T.D. Drivas, T.M. Elgindi, G. Iyer, and I.-J. Jeong. Anomalous dissipation in passive scalar transport. *Arch. Ration. Mech. Anal.* **243** (2022), no. 3, 1151–1180.
- [34] J. Duchon and R. Robert. Inertial energy dissipation for weak solutions of incompressible Euler and Navier-Stokes equations. *Nonlinearity* **13** (2000), no. 1, 249–255.
- [35] T.M. Elgindi and K. Liss. Norm growth, non-uniqueness, and anomalous dissipation in passive scalars. *Arch. Ration. Mech. Anal.* **248** (2024), no. 6:120.
- [36] G.L. Eyink. Energy dissipation without viscosity in ideal hydrodynamics I. Fourier analysis and local energy transfer. *Physica D: Nonlinear Phenomena* **78** (1994), no. 3-4, 222–240.
- [37] G.L. Eyink. Onsager’s “ideal turbulence” theory. *Journal of Fluid Mechanics* **988** (2024): P1.

- [38] G.L. Eyink and D. Benveniste. Suppression of particle dispersion by sweeping effects in synthetic turbulence. *Phys. Rev. E* **87** (2013), 023011.
- [39] G.L. Eyink and K.R. Sreenivasan. Onsager and the theory of hydrodynamic turbulence. *Reviews of modern physics* **78** (2006), no. 1, 87–135.
- [40] G. Falkovich, K. Gawedzki, and M. Vergassola. Particles and fields in fluid turbulence. *Reviews of modern Physics* **73** (2001), no. 4, 913–975.
- [41] A. Fannjiang and G. Papanicolaou. Convection enhanced diffusion for periodic flows. *SIAM J. Appl. Math.* **54** (1994), no. 2, 333–408.
- [42] U. Frisch. *Turbulence*. Cambridge University Press, Cambridge, 1995.
- [43] J. Geng and Z. Shen. Convergence rates in parabolic homogenization with time-dependent periodic coefficients. *J. Funct. Anal.* **272** (2017), no. 5, 2092–2113.
- [44] V. Giri, H. Kwon, M. Novack. The L^3 -based strong Onsager theorem. *Ann. of Math.* (2) (2025), to appear.
- [45] V. Giri and R.-O. Radu. The Onsager conjecture in 2D: a Newton-Nash iteration. *Invent. Math.* **238** (2024), no. 2, 691–768.
- [46] L. Huysmans and E.S. Titi. Non-uniqueness & inadmissibility of the vanishing viscosity limit of the passive scalar transport equation. *Journal de Mathématiques Pures et Appliquées* **198** (2025), 103685.
- [47] P. Isett. A proof of Onsager's conjecture. *Ann. of Math.* (2) **188** (2018), no. 3, 871–963.
- [48] V.V. Jikov, S.M. Kozlov and O.A Oleinik, *Homogenization of differential operators and integral functionals*. Springer-Verlag, Berlin, 1994.
- [49] C.J.P. Johansson and M. Sorella. Nontrivial absolutely continuous part of anomalous dissipation measures in time. 2023, ArXiV preprint 2303.09486.
- [50] C.J.P. Johansson and M. Sorella. Anomalous dissipation via spontaneous stochasticity with a two-dimensional autonomous velocity field. 2024, ArXiV preprint 2409.03599.
- [51] A. Kolmogorov. Local structure of turbulence in an incompressible fluid at very high reynolds number. *Dokl. Acad. Nauk SSSR* **30** (1941), no. 4, 299–303.
- [52] R.H. Kraichnan. Small-scale structure of a scalar field convected by turbulence. *Phys. Fluids* **11** (1968), no. 5, 945–953.
- [53] Y. Le Jan and O. Raimond. Integration of Brownian vector fields. *The Annals of Probability* **30** (2002), no. 2, 826–873.
- [54] A.J. Majda and P.R. Kramer. Simplified models for turbulent diffusion: theory, numerical modelling, and physical phenomena. *Phys. Rep.* **314** (1999), no. 4-5, 237–574.
- [55] A. L. Mazzucato, Y. Feng and C. Nobili. Enhanced dissipation by advection and applications to PDEs. 2025, ArXiV preprint 2501.17695.
- [56] D.W. McLaughlin, G.C. Papanicolaou, and O.R. Pironneau. Convection of microstructure and related problems. *SIAM Journal on Applied Mathematics* **45** (1985), no. 5, 780–797.
- [57] S. Modena and L. Székelyhidi Jr. Non-uniqueness for the transport equation with Sobolev vector fields. *Ann. PDE* **4** (2018), no. 2, Paper No. 18, 38 pp.
- [58] S. Modena and L. Székelyhidi Jr. Non-renormalized solutions to the continuity equation. *Calc. Var. Partial Differential Equations* **58** (2019), no. 6, Paper No. 208, 30 pp.
- [59] T. Moerschell and M. Sorella. Non-uniqueness of parabolic solutions for advection-diffusion equation. 2024, ArXiV preprint 2410.03308.
- [60] M. Novack and V. Vicol. An intermittent Onsager theorem. *Invent. Math.* **233** (2023), no. 1, 223–323.
- [61] A.M. Obukhov. The structure of the temperature field in a turbulent flow. *Izvestiya Akad. Nauk SSSR. Ser. Geograf. Geofiz.* **13** (1949), 58–69.
- [62] K. Rowan. On anomalous diffusion in the Kraichnan model and correlated-in-time variants. *Arch. Rational Mech. Anal.* **248** (2024), 93.
- [63] K. Rowan and E. Hess-Childs. A universal total anomalous dissipator. 2025, ArXiV preprint 2501.18526.
- [64] G. A. Pavliotis and A. M. Stuart. *Multiscale methods*, Texts in Applied Mathematics, 53, Springer, New York, 2008.
- [65] Z. Shen. *Periodic homogenization of elliptic systems*, Operator Theory: Advances and Applications Advances in Partial Differential Equations (Basel), 269, Birkhäuser/Springer, Cham, 2018.
- [66] B.I. Shraiman and E.D. Siggia. Scalar turbulence. *Nature* **405** (2000), no. 6787, 639–646.
- [67] K.R. Sreenivasan. The passive scalar spectrum and the Obukhov-Corrsin constant. *Phys. Fluids* **8** (1996), no. 1, 189–196.
- [68] 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** (2010), no. 1916, 1561–1577.
- [69] 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** (1953), no. 1137, 186–203.
- [70] 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.

COURANT INSTITUTE OF MATHEMATICAL SCIENCES, NEW YORK UNIVERSITY, NEW YORK, NY 10012, USA.

Email address: scotta@cims.nyu.edu

LABORATOIRE JACQUES-LOUIS LIONS, SORBONNE UNIVERSITY, PARIS, 75005, FRANCE.

Email address: scott.armstrong@sorbonne-universite.fr

COURANT INSTITUTE OF MATHEMATICAL SCIENCES, NEW YORK UNIVERSITY, NEW YORK, NY 10012, USA.

Email address: vicol@cims.nyu.edu