OPEN ACCESS
Mathematical Models and Methods in Applied Sciences
Vol. 31, No. 2 (2021) 223–292

DOI: 10.1142/S0218202521500068

Statistical solutions of the incompressible Euler equations

S. Lanthaler*, S. Mishra† and C. Parés-Pulido‡
Seminar for Applied Mathematics (SAM),
Eidgenössische Technische Hochschule Zürich (ETHZ),
Rämistrasse 101, 8092 Zürich, Switzerland
*samuel.lanthaler@sam.math.ethz.ch
†siddhartha.mishra@sam.math.ethz.ch
‡carlos.pares-pulido@sam.math.ethz.ch

Received 16 October 2019
Revised 22 October 2020
Accepted 17 November 2020
Published 16 February 2021
Communicated by F. Brezzi

We propose and study the framework of dissipative statistical solutions for the incompressible Euler equations. Statistical solutions are time-parameterized probability measures on the space of square-integrable functions, whose time-evolution is determined from the underlying Euler equations. We prove partial well-posedness results for dissipative statistical solutions and propose a Monte Carlo type algorithm, based on spectral viscosity spatial discretizations, to approximate them. Under verifiable hypotheses on the computations, we prove that the approximations converge to a statistical solution in a suitable topology. In particular, multi-point statistical quantities of interest converge on increasing resolution. We present several numerical experiments to illustrate the theory.

Keywords: Statistical solutions; incompressible Euler; Monte Carlo; structure functions; energy spectra.

AMS Subject Classification 2020: 35Q31, 76D06, 65M12, 65M30, 65M70, 65M75

1. Introduction

Many interesting incompressible fluid flows are characterized by high to very high values of the *Reynolds number*. Taking the infinite Reynolds number limit of the underlying Navier–Stokes equations (formally), results in the *Incompressible Euler* equations governing the motion of an ideal i.e. inviscid and incompressible fluid,

*Corresponding author.

given by

$$\begin{cases} \partial_t u + u \cdot \nabla u + \nabla p = 0, \\ \operatorname{div}(u) = 0, \\ u|_{t=0} = \overline{u}. \end{cases} \tag{1.1}$$

Here, the velocity field is denoted by $u \in \mathbb{R}^d$ (for $d = 2, 3$), and the pressure is denoted by $p \in \mathbb{R}_+$. The pressure acts as a Lagrange multiplier to enforce the divergence-free constraint.[38] The equations need to be supplemented with suitable boundary conditions. Throughout this paper, we shall assume periodic boundary conditions, and take as our domain D, the d-dimensional torus $D = \mathbb{T}^d$, $d \in \{2, 3\}$.

1.1. *Well-posedness results*

The question of global (in time) well-posedness of classical i.e. C^1 solutions of the incompressible Euler equations (1.1) in three space dimensions, even with sufficiently smooth initial data $\bar{u}$, is not yet resolved. Moreover in two space dimensions, where one can prove well-posedness of classical solutions as long as the initial data is sufficiently regular, many interesting initial data of interest, such as vortex sheets, do not possess this regularity. Hence, it is imperative to consider the so-called *weak solutions* of (1.1), defined as,

Definition 1.1. A vector field $u \in L^\infty([0,T); L^2(D; \mathbb{R}^d))$ is a weak solution of the incompressible Euler equations with initial data $\overline{u} \in L^2(D; \mathbb{R}^d)$, if

$$\int_0^T \int_D u \cdot \partial_t \phi + (u \otimes u) : \nabla \phi \, dx dt = - \int_D \overline{u} \cdot \phi(x, 0) dx, \tag{1.2}$$

for all test vector fields, $\phi \in C_c^\infty([0,T) \times D; \mathbb{R}^d)$, $\operatorname{div}(\phi) = 0$, and

$$\int_D u \cdot \nabla \psi \, dx = 0, \tag{1.3}$$

for all test functions $\psi \in C^\infty(D)$.

It is customary to require additional admissibility criteria in order to recover uniqueness of weak solutions. A natural criterion in this context is given by the so-called dissipative or admissible weak solutions, i.e. weak solutions u such that $\|u(t)\|_{L^2} \leq \|\overline{u}\|_{L^2}$. Although the global existence of admissible weak solutions in three space dimensions is open, one can prove global existence of admissible weak solutions in two dimensions with very general initial data. The most general result of this kind is the celebrated work of Delort,[10] later slightly extended to an edge case by Vecchi and Wu,[52] where global weak solutions of (1.1) were shown to exist as long as the initial data belongs to the so-called *Delort class*, i.e. with initial vorticity $\overline{\omega} \in H^{-1} \cap (\mathcal{M}_+ + L^1)$ the sum of a signed Radon measure (in space) and a L^1 function. In addition to being energy admissible, the constructed solutions satisfy the an additional *a priori* bound $\|\omega(t)\|_{\mathcal{M}} \leq \|\overline{\omega}\|_{\mathcal{M}}$. The existence of solutions with initial vorticity in the more natural space $H^{-1} \cap \mathcal{M}$ without any sign restriction remains an open problem. Uniqueness of weak solutions in the Delort class also remains open. On the other hand, it is well-known that – without the requirement of any bounds on the vorticity – energy admissible weak solutions are not necessarily unique.[9,43,44] Using convex integration techniques introduced by DeLellis and Szekelyhidi,[9] it has been shown that, in both two and three dimensions, there exists a dense set of "wild initial data" possessing infinitely many admissible weak solutions. Furthermore, it was shown in Ref. 49 that an explicit example of such wild initial data is given by the flat vortex sheet of constant vortex sheet strength in the two-dimensional case. This result has recently been extended to vortex sheet initial data along curves with sufficient Hölder regularity and not necessarily distinguished sign.[39] Vortex sheet initial data pose many open challenges in the two-dimensional case, and will thus be of particular interest in the current work.

1.2. *Numerical schemes*

A variety of numerical methods have been proposed in order to approximate the incompressible Euler equations (1.1). These include the spectral[21,41] (see also Ref. 7 and references therein for a general overview) and spectral viscosity methods,[50] that are particularly suitable for periodic boundary conditions. On bounded domains, efficient methods such as the finite difference projection method[3,8] and discontinuous Galerkin (DG) finite element method[22] have been proposed. Alternatives include numerically approximating the Euler equations in the vorticity-stream function formulation. Methods such as the Lagrangian vortex blob,[34] point vortex[23] and central finite difference schemes[30] have been proposed in this context.

Rigorous convergence results for numerical approximations of the incompressible Euler equations (1.1) are mostly restricted to two space dimensions and to sufficiently regular initial data, see Ref. 2 for spectral viscosity methods, Ref. 38 for vortex methods and Ref. 22 for DG methods. Notable exceptions include Ref. 36 where the central schemes of Ref. 30 were shown to converge to weak solutions of the two-dimensional Euler equations as long as the initial vorticity was in L^p, for $p > 1$. Similarly in Refs. 34 and 35, the authors showed convergence of vortex point and vortex blob methods in two space dimensions as long as the initial vorticity was a signed Radon measure. Finally in a recent paper,[26] the authors showed convergence of a spectral viscosity method for initial data in the Delort class.

Furthermore, careful numerical experiments, for instance those presented in Refs. 25 and 17 have shown that numerical schemes may not necessarily converge, even in two space dimensions, for rough initial data. Even when the numerical approximations converge (such as in Ref. 26), the inherent instabilities of the Euler equations lead to a very slow convergence rate (see also Fig. 8(left)) and render the computation of weak solutions of (1.1) prohibitively expensive.

1.3. *Measure-valued and statistical solutions*

Given the lack of well-posedness results for weak solutions and the lack of convergent numerical approximations, there is considerable scope for the design of alternative solution frameworks for (1.1). One such framework is that of measure-valued solutions,[11] where the sought for solutions are no longer functions but space-time parameterized probability measures on state space. The global existence of measure-valued solutions, even in three space dimensions, was shown in Ref. 11. A convergent numerical method (of the spectral viscosity type) and an efficient algorithm to compute measure-valued solutions was proposed in Ref. 25. The convergence to the limiting Young measure has conventionally been considered in the weak-$*$ topology. More recently, a slightly stronger notion of convergence, termed $\mathcal{K}$-convergence, has been introduced.[12] It has been shown[12] that (up to a subsequence) Césaro-type means of approximate solutions computed at different numerical resolutions converge pointwise almost everywhere to the limiting measure-valued solution.

However, measure-valued solutions are generically non-unique. This holds true even for the much simpler case of the one-dimensional Burgers' equation.[45] In Ref. 14, the authors implicated the lack of information about multi-point (spatial) correlations in the non-uniqueness of measure-valued solutions. Moreover, they also proposed a framework of *statistical solutions* as an attempt to recover uniqueness.

In the formulation of Ref. 14, statistical solutions are time-parameterized probability measures on L^p, for $1 \leq p < \infty$, that are consistent with the underlying PDE in a weak sense. They were shown to be equivalent to a family of *correlation measures*, with the kth member of this family is a Young measure representing correlations (or joint probabilities) of the solution at k distinct spatial points. Thus, one can interpret statistical solutions as measure-valued solutions, augmented with information about all possible multi-point correlations. The consideration of multi-point statistics is one of the main differences of the present work with earlier contributions such as Ref. 25, which focused on the computation of a measure-valued solution, i.e. single-point statistics. *A priori*, statistical solutions contain much more information than measure-valued solutions. Moreover, statistical solutions encode statistical (ensemble averaged) properties of the solutions of the underlying PDE. Thus, statistical solutions provide a suitable framework for uncertainty quantification (UQ).[1,14] This is particularly relevant for the incompressible Euler equations as it is well-known that the flow of fluids at very high Reynolds numbers can be turbulent, and only (or statistical) properties can be inferred from measurements.[20]

Statistical solutions for scalar conservation laws were considered in Ref. 14, wherein well-posedness was shown under an entropy condition. In particular, information about infinitely many correlations was necessary to ensure uniqueness. In Refs. 15 and 16, a Monte Carlo algorithm, based on the ensemble averaging algorithm of Ref. 13, was proposed and analyzed for scalar conservation laws and multi-dimensional hyperbolic systems of conservation laws, respectively. In contrast to Ref. 16 where multi-dimensional hyperbolic systems of conservation laws were considered, we focus on the case of incompressible Euler equations in this paper. Although the concept of statistical solutions is similar in both cases, there are important differences that we highlight in this paper. In particular, the theory for deterministic solutions of the incompressible Euler equations is much more developed than its counterpart for systems of conservation laws. This will allow us to prove global existence and uniqueness of dissipative statistical solutions for the two-dimensional Euler equations, and convergence of our approximate statistical solutions, whenever the initial measure $\overline{\mu}$ is concentrated on sufficiently smooth data, without any additional assumptions on the structure functions. Moreover, we will focus on inferring information about multi-point correlations from energy spectra. This approach differs from Ref. 16, where the focus was entirely on structure functions.

Independent notions of statistical solutions of the incompressible Navier–Stokes equations have been proposed in Refs. 19 and 54. While the statistical solutions of Foias, Rosa and Temam[19] are formulated in terms of the evolution equations of integrals of functionals $\int_H \Phi(u)\, d\mu_t(u)$ on a suitable Hilbert space H, the statistical solutions in the present work are formulated in terms of an infinite family of PDEs for the multi-point correlation measures $\nu^k_{t,x_1,\dots,x_k}(\xi_1,\dots,\xi_k)$. These correlation measures encode the probability of the flow field $u(t,x)$ attaining certain values at points $x_1,\dots,x_k$ and time t, i.e. one might informally write

$$\nu^k_{t,x_1,\dots,x_k}(\xi_1,\dots,\xi_k) = \mathrm{Prob}[u(t,x_1)=\xi_1,\dots,u(t,x_k)=\xi_k].$$

Despite the apparent differences between the current work and Ref. 19, the two approaches can be related to each other, using the correspondence between multi-point correlation measures and infinite-dimensional measures established in Ref. 14. The precise link between the zero-viscosity limit of the statistical solutions of the Navier–Stokes equations of Foias, Rosa and Temam, and the statistical solutions of the present work will be considered in a forthcoming paper.

1.4. *Aims and scope*

The main goal of this paper is to propose a suitable notion of statistical solutions for the incompressible Euler equations (1.1) and to analyze and approximate these statistical solutions. To this end, we will

- propose a notion of dissipative statistical solutions, i.e. a time-parameterized probability measure on $L^p(D)$ that is consistent with (1.1) in a suitable sense and prove partial well-posedness in some special cases, namely local in time well-posedness and global well-posedness for sufficiently regular initial data in two space dimensions,
- propose a numerical algorithm, based on ensemble averaging and a spectral viscosity spatial discretization, to approximate statistical solutions of (1.1) and prove that the approximations converge in an appropriate topology to a statistical solution, under reasonable and verifiable hypotheses on the numerical method,
- perform and present extensive numerical experiments to verify the theory and to illustrate interesting properties of statistical solutions.

The rest of the paper is organized as follows: in Sec. 2, we present time-parameterized probability measures on $L^2(D;\mathbb{R}^d)$ and characterize convergence in a suitable topology on this space of measures. In Sec. 3, we define statistical solutions of (1.1) and present partial well-posedness results. The numerical approximation of statistical solutions and its convergence is presented in Sec. 4 and numerical experiments are presented in Sec. 5.

2. Time-Parameterized Probability Measures on $L^2(D;\mathbb{R}^d)$

As mentioned in the introduction, statistical solutions are time-parameterized probability measures on L^p. For incompressible Euler equations, the energy bound enforces that $p = 2$. In this section, we will describe time-parameterized probability measures, characterize them and describe a suitable topology on them. Although different in several details, similar considerations have previously appeared in different contexts in Refs. 14 and 16. To streamline our discussion and to avoid distracting the reader with technical details, we will therefore merely state the core results in this section. For completeness, detailed proofs of these results have been included in the appendices.

2.1. *Notation*

For spatial dimensions $d = 2, 3$, we denote the spatial domain $D = \mathbb{T}^d$, i.e. d-dimensional torus $[-\pi, \pi]^d$, with endpoints identified, the co-domain (or phase space) $U = \mathbb{R}^d$, so that a solution of (1.1) is given as a vector field $u : D \times [0, T) \to U$. On occasion $x \mapsto u(x, t)$ will also be identified with a 2π-periodic function on $\mathbb{R}^d$ in the obvious way.

As we are interested in time-dependent vector fields $u(x, t)$, we will fix a final time $T > 0$ and consider $u(x, t)$ as a mapping $u : [0, T) \to L^2_x = L^2(D; U)$. In general, for a Banach space Y, we will denote by $L^q_t(Y)$ the space of measurable functions $u : [0, T) \to Y$, such that

$$\|u\|_{L^q_t(Y)} = \left(\int_0^T \|u\|^q_Y dt \right)^{1/q} < \infty.$$

Of particular interest in the context of the incompressible Euler equations are spaces $L^q_t L^2_x = L^q_t(L^2_x)$, for $1 \le q \le \infty$.

For a function $u \in L^2_x$, we will denote by $\widehat{u}(k)$ the kth Fourier coefficient ($k \in \mathbb{Z}^d$), i.e.

$$\widehat{u}(k) = \frac{1}{(2\pi)^d} \int_D u(x) e^{-ikx} dx \Rightarrow u(x) = \sum_{k \in \mathbb{Z}^d} \widehat{u}(k) e^{ikx}.$$

We will say that $\omega(r)$ is a *modulus of continuity* if $\omega : [0, \infty) \to [0, \infty)$ is a map such that $\lim_{r \to 0} \omega(r) = 0$.

On multiple occasions in this paper, we will need to mollify a given function $x \mapsto u(x)$. Let us therefore fix a standard mollifier $\rho(x) \in C^\infty(\mathbb{R}^d)$, supported in a ball of radius 1 around the origin, i.e. $\rho \geq 0$ is a function, such that

$$\int_{\mathbb{R}^d} \rho(x)dx = 1, \quad \rho(x) = 0, \quad \text{if } |x| > 1.$$

Given $\epsilon > 0$, we denote by $\rho_\epsilon \in C_c^\infty(\mathbb{R}^d)$ the function $\rho_\epsilon(x) = \epsilon^{-d}\rho(x/\epsilon)$. The *mollification* of $u(x)$ is now defined by convolution with $\rho_\epsilon(x)$, as

$$u_\epsilon(x) := (u * \rho_\epsilon)(x) = \int_{\mathbb{R}^d} u(x-y)\rho_\epsilon(y)dy.$$

2.2. *Weak convergence of probability measures*

Given a topological space X, we denote by $\mathcal{P}(X)$ the space of all probability measures on the Borel σ-algebra of X. Given a sequence $\mu_n \in \mathcal{P}(X)$ $(n \in \mathbb{N})$ and $\mu \in \mathcal{P}(X)$, we say that μ_n *converges weakly* to μ, written $\mu_n \rightharpoonup \mu$, if

$$\int_X F(u)d\mu_n(u) \to \int_X F(u)d\mu(u), \tag{2.1}$$

for all $F \in C_b(X)$, where $C_b(X)$ is the space of bounded, continuous functions on X.

We shall call a family of probability measures $\{\mu^\Delta\}_{\Delta>0} \subset \mathcal{P}(X)$ defined on a separable Banach space X *tight*, provided that for any $\epsilon > 0$ there exists a compact subset $K \subset X$ such that

$$\mu^\Delta(K) \geq 1 - \epsilon, \quad \forall\, \Delta > 0.$$

It is a classical result due to Prokhorov (see e.g. Theorems 8.6.7, 8.6.8 of the monograph[4]) that a family $\mu^\Delta \in \mathcal{P}(X)$, with X a separable Banach space, is tight if and only if μ^Δ is relatively compact under the weak topology.

We recall that weak convergence on a Banach space X is metrized by the p-*Wasserstein metric* W_p:

$$W_p(\mu,\nu) := \left(\inf_{\pi\in\Pi(\mu,\nu)} \int_{X\times X} \|u-v\|_X^p \, d\pi(u,v) \right)^{1/p}, \tag{2.2}$$

where the infimum is taken over all *transfer plans* $\pi \in \Pi$ defined as

$$\Pi(\mu,\nu) := \left\{ \pi \in \mathcal{P}(X \times X) : \int_{X\times X} (F(u)+G(v))d\pi(u,v) \right.$$
$$\left. = \int_X F(u)d\mu(u) + \int_X G(v)d\nu(v) \right\},$$

for all measurable F, G. More precisely, let μ^Δ be a sequence of probability measures on X. If there exists a $M > 0$, such that $\mu^\Delta \in \mathcal{P}(X)$ are uniformly concentrated on the ball $B_M(0) = \{u \in X \,|\, \|u\|_X < M\}$ of radius M and centered at the origin in X, then

$$\mu^\Delta \rightharpoonup \mu \Leftrightarrow W_p(\mu^\Delta, \mu) \to 0,$$

where $1 \le p < \infty$. In this paper, we shall only consider $p = 1$ and $p = 2$. The space of probability measures μ on which W_p is well-defined, i.e. such that $\int_X \|u\|^p \, d\mu(u) < \infty$ is usually denoted by $\mathcal{P}_p(X)$. It is well-known that $(\mathcal{P}_p, W_p)$ is a complete metric space.[53]

We also note that for $p = 1$, the following *duality formula* holds:

$$W_1(\mu, \nu) = \sup_{\Phi} \int_X \Phi(u)[d\mu(u) - d\nu(u)], \tag{2.3}$$

where the supremum is taken over all 1-Lipschitz continuous functions $\Phi : X \to \mathbb{R}$.

Next, we characterize the compactness properties of families of probability measures $\mu^\Delta \in \mathcal{P}(L^2_x)$. The following quantity, the so-called *structure function*, is of particular interest in the present context:

$$S^2_r(\mu) := \left(\int_{L^2_x} \int_D \fint_{B_r(0)} |u(x+h) - u(x)|^2 dh dx d\mu(u) \right)^{1/2},$$

where $B_r(0) = \{\xi \in \mathbb{R}^d \,|\, |\xi| < r\}$, and $\fint_{B_r(0)} := \frac{1}{|B_r(0)|} \int_{B_r(0)}$ denotes the mean over $B_r(0)$.

We now state our main result characterizing certain compact subsets of $\mathcal{P}(L^2_x)$.

Theorem 2.1. *Let $\mathcal{F} \subset \mathcal{P}(L^2_x)$ be a family of probability measures on L^2_x. Assume that there exists $M > 0$, such that $\mu(B_M(0)) = 1$ for all $\mu \in \mathcal{F}$, where $B_M(0) = \{u \in L^2_x \,|\, \|u\|_{L^2_x} < M\}$. Then the following statements are equivalent*:

(i) *$\mathcal{F} \subset L^2_x$ has compact closure (with respect to the weak topology)*,
(ii) *There exists a modulus of continuity ω, such that we have a uniform bound on the structure function*:

$$S^2_r(\mu) \le \omega(r), \quad \forall \mu \in \mathcal{F}.$$

The proof of this theorem can be found in Appendix A.

Remark 2.1. Theorem 2.1 is closely related to Kolmogorov's characterization of compact subset of L^2_x: Indeed, there is a natural isometric embedding

$$(L^2_x, \|\cdot\|_{L^2_x}) \hookrightarrow (\mathcal{P}_1(L^2_x), W_1), \quad u \mapsto \delta_u.$$

Thus, a bounded set $K \subset L^2_x$ has compact closure if, and only if, its image under this embedding $\{\delta_u \,|\, u \in K\} \subset \mathcal{P}(L^2_x)$ has compact closure. Using Theorem 2.1, we conclude that a bounded set $K \subset L^2_x$ has compact closure if, and only if, satisfies Kolmogorov's equicontinuity property, which corresponds to property (2) in Theorem 2.1.

2.3. *Time parameterized probability measures*

As mentioned before, statistical solutions are time-parameterized probability measures. We define them in what follows.

Definition 2.1. We denote by $L^1_t(\mathcal{P}) = L^1([0,T);\mathcal{P})$ the space of weak-$*$ measurable mappings $[0,T) \to \mathcal{P}(L^2_x)$, namely mappings $t \mapsto \mu_t$ such $t \mapsto \int_{L^2_x} F(u)d\mu_t(u)$ is measurable for a.e. $t \in [0,T)$, for all $F \in C_b(L^2_x)$ and with the property that

$$\int_0^T \int_{L^2_x} \|u\|_{L^2_x} \, d\mu_t(u) \, dt < \infty.$$

Denoting by δ_0 the Dirac measure concentrated on $0 \in L^2_x$, the above condition can equivalently be written as

$$\int_0^T W_1(\delta_0, \mu_t) \, dt < \infty.$$

This leads us to define a natural metric on $L^1([0,T);\mathcal{P})$ by

$$d_T(\mu_t, \nu_t) := \int_0^T W_1(\mu_t, \nu_t) \, dt. \tag{2.4}$$

We then have the following proposition, whose proof is presented in Appendix B.

Proposition 2.1. *The metric space* $(L^1_t(\mathcal{P}), d_T)$ *is a complete metric space.*

Our objective is to find a suitable topology on $L^1_t(\mathcal{P})$ and characterize compactness in this topology. It would be natural to try and extend the compactness Theorem 2.1 to time-parameterized probability measures and find a suitable version of the weak topology. This necessitates formalizing some notion of time-continuity or time-regularity of underlying functions.

To this end, fix a (time-independent) divergence-free test function $\phi \in C_c^\infty(D;\mathbb{R}^d)$. Formally, solutions of the incompressible Euler equations (1.1) satisfy for $s,t \in [0,T)$,

$$\int_D [u(x,t) - u(x,s)]\phi(x) \, dx = \int_s^t \int_D u(x,\tau) \otimes u(x,\tau) : \nabla\phi(x) \, dx \, d\tau,$$

so that

$$\left| \int_D [u(x,t) - u(x,s)]\phi(x) \, dx \right| \le C \|u\|_{L^\infty_t L^2_x} \|\nabla\phi\|_{L^\infty_x} |t-s|.$$

Furthermore, we have a natural energy bound $\|u\|_{L^\infty_t L^2_x} \le \|u_0\|_{L^2_x}$, in terms of the initial data u_0. If $L > 0$ is large enough such that by Sobolev embedding $H^L_x = H^L(D;U) \hookrightarrow C^1(D;U)$, then it follows that

$$\left| \int_D [u(x,t) - u(x,s)]\phi(x) \, dx \right| \le C \|\phi\|_{H^L_x} |t-s|,$$

where the constant $C > 0$ depends only on the initial data. Taking the supremum over all $\phi \in H^L_x$ with $\|\phi\|_{H^L_x} \le 1$, it follows, at least formally, that

$$\|u(t) - u(s)\|_{H^{-L}_x} \le C|t-s|, \quad \forall\, s,t \in [0,T). \tag{2.5}$$

Given these considerations, it is natural to assume that statistical solutions of the Euler equations satisfy some version of this time continuity. A formalization is provided in the following definition,

Definition 2.2. A weak-$*$ measurable, time-parameterized probability measure $t \mapsto \mu_t \in \mathcal{P}(L^2_x)$ is called *time-regular*, if there exists a constant $L > 0$, and a mapping $s, t \mapsto \pi_{s,t} \in \mathcal{P}(L^2_x \times L^2_x)$, such that for almost all $s, t \in [0, T)$:

- The measure $\pi_{s,t}$ is a transport plan from μ_s to μ_t,
- There exists a constant $C > 0$, such that $\pi_{s,t}$ satisfies the following regularity condition

$$\int_{L^2_x \times L^2_x} \|u - v\|_{H^{-L}_x} \, d\pi_{s,t}(u, v) \leq C|t - s|.$$

A family $\{\mu^\Delta_t\}_{\Delta>0}$ of time-parameterized probability measures is *uniformly time-regular*, provided that each μ^Δ_t is time-regular, and the constants $L, C > 0$ above can be chosen independently of $\Delta > 0$.

Remark 2.2. Note that if μ_t is of the form

$$\mu_t = \frac{1}{M} \sum_{i=1}^{M} \delta_{u_i(t)},$$

with $t \mapsto u(t)$ weak solutions of the incompressible Euler equations satisfying (2.5), then we can define suitable transfer plans

$$\pi_{s,t} = \frac{1}{M} \sum_{i=1}^{M} \delta_{u_i(s)} \otimes \delta_{u_i(t)}.$$

The time-regularity property follows from the estimate (2.5) for the u_i.

We now show that a family μ^Δ_t, $\Delta > 0$, of uniformly time-regular probability measures is relatively compact, provided that they satisfy a time-averaged version of the second property of Theorem 2.1. To this end, we define the time-averaged structure function of $(t \mapsto \mu_t) \in L^1_t(\mathcal{P})$ (weak-$*$ measurable) as the following quantity:

$$S^2_r(\mu_t, T) := \left(\int_0^T \int_{L^2_x} \int_D \fint_{B_r(0)} |u(x + h) - u(x)|^2 \, dh \, dx \, d\mu_t(u) \, dt \right)^{1/2}. \tag{2.6}$$

The main result of the present section is the following compactness result:

Theorem 2.2. *Let $\mu^\Delta_t \in L^1_t(\mathcal{P})$ be a family of uniformly time-regular probability measures, for $\Delta > 0$, for which there exists $M > 0$, such that $\mu^\Delta_t(B_M(0)) = 1$ for all $\Delta > 0$, a.e. $t \in [0, T)$. Here $B_M(0) := \{\|u\|_{L^2_x} < M\}$. If there exists a modulus of continuity $\omega(r)$ such that*

$$S^2_r(\mu^\Delta_t, T) \leq \omega(r), \quad \forall\, \Delta > 0,$$

then μ^Δ_t is relatively compact in $L^1_t(\mathcal{P})$.

The idea behind Theorem 2.2 is to use the spatial regularity of the sequence to show that the weak time-regularity assumption of Definition 2.2 implies a similar time-regularity with respect to a stronger spatial norm, where H^{-L} is replaced by L^2. The details of this argument are provided in Appendix C.

Let us also remark that a limit $\mu_t^\Delta \to \mu_t$ of a uniformly time-regular sequence μ_t^Δ is itself time-regular (see Appendix D for a proof):

Proposition 2.2. *Let $\mu_t^\Delta \in L_t^1(\mathcal{P})$ be a family of uniformly time-regular probability measures, for $\Delta > 0$. And such that there exists $M > 0$ with $\mu_t^\Delta(B_M(0)) = 1$ for all $\Delta > 0$, a.e. $t \in [0, T)$, where $B_M(0) := \{\|u\|_{L_x^2} < M\}$. If $\mu_t^\Delta \to \mu_t$ in $L_t^1(\mathcal{P})$, then μ_t is time-regular in the sense of Definition* 2.2, *with the same time-regularity constants $C, L > 0$ as for the family μ_t^Δ.*

2.4. *Time-dependent correlation measures and their compactness*

It has been shown in Ref. 14 that there is a one-to-one correspondence between probability measures on L_x^2 and so-called correlation measures. Correlation measures are defined as infinite hierarchies of Young measures, taking into account spatial correlations, or more precisely,

Definition 2.3. A *correlation measure* is a collection $\boldsymbol{\nu} = (\nu^1, \nu^2, \dots)$ of maps $\nu^k : D^k \to \mathcal{P}(U^k)$ satisfying the following properties:

(1) *Weak-$*$ measurability*: Each map $\nu^k : D^k \to \mathcal{P}(U^k)$ is weak-$*$-measurable, in the sense that the map $x \mapsto \langle \nu_x^k, f\rangle$ from $x \in D^k$ into $\mathbb{R}$ is Borel measurable for all $f \in C_0(U^k)$ and $k \in \mathbb{N}$. In other words, ν^k is a Young measure from D^k to U^k.
(2) *L^2-boundedness*: $\boldsymbol{\nu}$ is L^2-bounded, in the sense that

$$\int_D \langle \nu_x^1, |\xi|^2\rangle \, dx < +\infty. \tag{2.7}$$

(3) *Symmetry*: If σ is a permutation of $\{1, \dots, k\}$ and $f \in C_0(\mathbb{R}^k)$ then $\langle \nu_{\sigma(x)}^k, f(\sigma(\xi))\rangle = \langle \nu_x^k, f(\xi)\rangle$ for a.e. $x \in D^k$. Here, we denote $\sigma(x) = \sigma(x_1, x_2, \dots, x_k) = (x_{\sigma_1}, x_{\sigma_2}, \dots, x_{\sigma_k})$. $\sigma(\xi)$ is denoted analogously.
(4) *Consistency*: If $f \in C_0(U^k)$ is of the form $f(\xi_1, \dots, \xi_k) = g(\xi_1, \dots, \xi_{k-1})$ for some $g \in C_0(U^{k-1})$, then $\langle \nu_{x_1,\dots,x_k}^k, f\rangle = \langle \nu_{x_1,\dots,x_{k-1}}^{k-1}, g\rangle$ for almost every $(x_1, \dots, x_k) \in D^k$.
(5) *Diagonal continuity* (*DC*): If $B_r(x) := \{y \in D : |x - y| < r\}$ then

$$\lim_{r\to 0} \int_D \fint_{B_r(x)} \langle \nu_{x,y}^2, |\xi_1 - \xi_2|^2\rangle \, dy \, dx = 0. \tag{2.8}$$

Each element ν^k is called a *correlation marginal*. We let $\mathfrak{L}^2 = \mathfrak{L}^2(D, U)$ denote the set of all correlation measures from D to U.

It has been shown in Ref. 14, that if $\mu \in \mathcal{P}(L^2_x)$, then μ is associated with a unique correlation measure $\boldsymbol{\nu}$, with the interpretation that for $A_1, \dots, A_k \subset U$:

$$\mu[u(x_i) \in A_i, i = 1, \dots, k] = \nu^k_{x_1,\dots,x_k}(A_1 \times \cdots \times A_k).$$

More precisely, we have the following theorem[14]:

Theorem 2.3. *For every correlation measure* $\boldsymbol{\nu} \in \mathfrak{L}^2(D, U)$ *there exists a unique probability measure* $\mu \in \mathcal{P}(L^2(D; U))$ *satisfying*

$$\int_{L^2_x} \|u\|^2_{L^2_x} \, d\mu(u) < \infty, \tag{2.9}$$

such that

$$\int_{D^k} \int_{U^k} g(x, \xi) \, d\nu^k_x(\xi) \, dx = \int_{L^2_x} \int_{D^k} g(x, u(x)) \, dx \, d\mu(u), \tag{2.10}$$

for all $g \in C_0(D^k \times U^k)$ *and* $k \in \mathbb{N}$ *(where* $u(x)$ *denotes the vector* $(u(x_1), \dots, u(x_k))$*). Conversely, for every probability measure* $\mu \in \mathcal{P}(L^2(D; U))$ *with finite moment* (2.9), *there exists a unique correlation measure* $\boldsymbol{\nu} \in \mathfrak{L}^2(D, U)$ *satisfying* (2.10). *The relation* (2.10) *is also valid for any measurable* $g : D \times U \to \mathbb{R}$ *such that* $|g(x, \xi)| \leq C|\xi|^2$ *for a.e.* $x \in D$.

Moreover, the moments

$$m^k : D^k \mapsto U^{\otimes k}, \quad m^k(x) = \langle \nu^k_x, \xi_1 \otimes \xi_2 \otimes \cdots \otimes \xi_k \rangle, \tag{2.11}$$

uniquely determine the correlation measure ν *and hence the underlying probability measure* μ.

The following result is obtained as a consequence of Theorem 2.2 (for a proof, see Appendix E):

Theorem 2.4. *Let* $\{\mu^\Delta_t\}_{\Delta > 0}$ *be a family of uniformly time-regular probability measures in* $L^1_t(\mathcal{P})$, *and assume that there exists* $M > 0$, *such that* $\mu^\Delta_t(B_M) = 1$ *for all* $\Delta > 0$ *and* $t \in [0, T)$. *Let* $\boldsymbol{\nu}^\Delta_t = (\nu^{\Delta,1}_t, \nu^{\Delta,2}_t, \dots)$ *denote the corresponding time-parameterized correlation measures. If there exists a uniform modulus of continuity* $\omega(r)$, *such that*

$$\int_0^T \int_D \fint_{B_r(x)} \langle \nu^{\Delta,2}_{t,x,y}, |\xi_1 - \xi_2|^2 \rangle \, dy \, dx \, dt \leq \omega(r), \quad \forall \, \Delta > 0,$$

then $\{\mu^\Delta_t\}_{\Delta > 0}$ *is relatively compact in* $L^1_t(\mathcal{P})$, *i.e. there exists a subsequence* $\Delta_j \to 0$ $(j \in \mathbb{N})$, *and a time-parameterized probability measure* $\mu_t \in L^1_t(\mathcal{P})$, *such that*

$$\int_0^T W_1(\mu^{\Delta_j}_t, \mu_t) \, dt \to 0, \quad \text{as } j \to \infty.$$

Furthermore, denoting by $\boldsymbol{\nu}_t = (\nu^1_t, \nu^2_t, \dots)$ *the correlation measure corresponding to the limit* μ, *we have*

- L^2-*bound*: $\int_D \langle \nu^1_{t,x}, |\xi|^2 \rangle \, dx \leq M^2$, *for a.e.* $t \in [0, T)$,

- *the two-point correlations satisfy*

$$\int_0^T \int_D \fint_{B_r(x)} \langle \nu^2_{t,x,y}, |\xi_1 - \xi_2|^2 \rangle \, dy \, dx \, dt \le \omega(r).$$

- *We define admissible observables, in terms of test functions* $g \in C([0,T) \times D^k \times U^k)$, *which satisfy the following bounds*:

$$\begin{aligned} |g(t,x,\xi)| &\le C \prod_{i=1}^k (1 + |\xi_i|^2), \\ |g(t,x,\xi) - g(t,x,\xi')| &\le C \sum_{i=1}^k \Pi_i(\xi,\xi') \sqrt{1 + |\xi_i|^2 + |\xi_i'|^2} |\xi_i - \xi_i'|, \end{aligned} \tag{2.12}$$

where $C > 0$ *is a fixed constant, independent of* $t \in [0,T)$, $x \in D^k$ *and* $\xi, \xi' \in U^k$. *Here* $\Pi_i(\xi,\xi')$ *is defined as*

$$\Pi_i(\xi,\xi') := \prod_{\substack{j=1 \\ j \ne i}}^k (1 + |\xi_j|^2 + |\xi_j'|^2), \quad \xi, \xi' \in U^k. \tag{2.13}$$

Then, these admissible observables converge strongly in $L^1_{t,x}$, *in the sense that*

$$\lim_{j \to \infty} \int_0^T \int_{D^k} |\langle \nu^{\Delta_j,k}_{t,x}, g(x,\xi) \rangle - \langle \nu^k_{t,x}, g(x,\xi) \rangle| \, dx \, dt = 0.$$

A particular point of interest in the statement of previous theorem is the characterization of a suitable set of "admissible observables", whose convergence is assured by the convergence $\mu^\Delta_t \to \mu_t$ in $L^1_t(\mathcal{P})$.

Remark 2.3. We note that the uniform modulus of continuity estimate in Theorem 2.4 can equivalently be expressed as

$$\int_0^T \int_{L^2_x} \int_D \fint_{B_r(0)} |u(x+h) - u(x)|^2 \, dh \, dx \, d\mu^\Delta_t(u) \, dt \le \omega(r),$$

or $S^2_r(\mu^\Delta_t, T)^2 \le \omega(r)$, for all $\Delta > 0$.

3. Dissipative Statistical Solutions and their Well-Posedness

Given the discussion on time-parameterized probability measures in the last section, we can now define statistical solutions of (1.1) as the following.

Definition 3.1. A time-parameterized probability measure $\mu_t \in L^1_t(\mathcal{P})$ is a *statistical solution* of the incompressible Euler equations with initial data $\bar{\mu}$, if $t \mapsto \mu_t$ is time-regular, and the associated correlation measure $\boldsymbol{\nu}_t$ satisfies:

(1) Given $\phi_1, \ldots, \phi_k \in C^\infty([0,T) \times D; \mathbb{R}^d)$ with $\operatorname{div}(\phi_i) = 0$ for all $i = 1, \ldots, k$, set

$$\phi(t,x) = \phi_1(t,x_1) \otimes \cdots \otimes \phi_k(t,x_k), \quad \text{where } x = (x_1, \ldots, x_k).$$

Let us denote $F(\xi) := \xi \otimes \xi$ and define a contraction by

$$(\xi_1 \otimes \cdots \otimes F(\xi_i) \otimes \cdots \otimes \xi_k) : \nabla_{x_i}\phi = \left[\prod_{j\neq i}(\xi_j \cdot \phi_j)\right](\xi_i \cdot \nabla_{x_i}\phi_i)\cdot \xi_i.$$

Then $\nu^k = \nu^k_{t,x_1,\dots,x_k}$ satisfies

$$\int_0^T \int_{D^k} \Big\{ \langle \nu^k, \xi_1 \otimes \cdots \otimes \xi_k\rangle : \partial_t \phi + \sum_i \langle \nu^k, \xi_1 \otimes \cdots \otimes F(\xi_i) \otimes \cdots \otimes \xi_k\rangle : \nabla_{x_i}\phi \Big\}\, dx\, dt + \int_{D^k} \langle \bar{\nu}^k, \xi_1 \otimes \cdots \otimes \xi_k\rangle : \phi(0,x)\, dx = 0.$$

Here $\bar{\nu}$ is the correlation measure corresponding to the initial data $\bar{\mu}$.

(2) For all $\psi \in C_c^\infty(D)$, we have

$$\int_{D^2} \langle \nu^2_{t,x_1,x_2}, \xi_1 \otimes \xi_2\rangle : (\nabla\psi(x_1) \otimes \nabla\psi(x_2))\, dx_1\, dx_2 = 0,$$

for a.e. $t \in [0,T)$.

The above PDEs specify the time-evolution of the moments (2.11) for all k and by Theorem 2.3, determine the evolution of the probability measure μ_t.

Remark 3.1. As ν^1 above is a standard Young measure, it is straightforward to observe that the corresponding identity for the evolution of ν^1 corresponds to the definition of measure-valued solution of (1.1) in the sense of Ref. 11 under the further assumption that there is no concentration. Hence, one can think of statistical solutions as measure-valued solutions coupled with information about all possible multi-point correlations.

We first show that the second property of Definition 3.1 is equivalent to the requirement that μ_t be supported on divergence-free vector fields for almost all t.

Lemma 3.1. *Let $\mu \in \mathcal{P}(L^2_x)$, with associated correlation measure $\boldsymbol{\nu}$. Then μ is concentrated on divergence-free vector fields if, and only if,*

$$\int_{D^2} \langle \nu^2_{x_1,x_2}, \xi_1 \otimes \xi_2\rangle : (\nabla\psi(x_1) \otimes \nabla\psi(x_2))\, dx_1\, dx_2 = 0,$$

for all $\psi \in C_c^\infty(D)$.

Proof. Let $\psi \in C_c^\infty(D)$. Then, we have the following identity

$$\int_{L^2_x} \left[\int_D u \cdot \nabla\psi\, dx\right]^2 d\mu(u)$$

$$
\begin{aligned}
&= \int_{L^2_x} \int_D (u(x_1) \otimes u(x_2)) : (\nabla\psi(x_1) \otimes \nabla\psi(x_2))\, dx_1\, dx_2\, d\mu(u) \\
&= \int_D \langle \nu^2_{x_1,x_2}, \xi_1 \otimes \xi_2 \rangle : (\nabla\psi(x_1) \otimes \nabla\psi(x_2))\, dx_1\, dx_2. \qquad (3.1)
\end{aligned}
$$

Therefore, if μ is concentrated on the divergence-free vector fields, then

$$
\int_D u \cdot \nabla\psi\, dx = 0, \quad \mu\text{-almost surely}
$$

and hence, from Eq. (3.1), we obtain

$$
\int_D \langle \nu^2_{x_1,x_2}, \xi_1 \otimes \xi_2 \rangle : (\nabla\psi(x_1) \otimes \nabla\psi(x_2))\, dx_1\, dx_2 = 0. \qquad (3.2)
$$

To prove the converse, let us assume that relation (3.2) holds for all $\psi \in C^\infty_c$. Let $\psi_n \in C^\infty_c$, $n \in \mathbb{N}$, be a countable, dense subset of $H^1(D)$. Then, we have

$$
\{u \in L^2_x \,|\, \mathrm{div}(u) \neq 0 \text{ distributionally}\} = \bigcup_{n\in\mathbb{N}} \left\{ u \in L^2_x \,\middle|\, \int_D u \cdot \nabla\psi_n\, dx \neq 0 \right\}.
$$

Set now $F_n(u) := [\int_D u \cdot \nabla\psi_n\, dx]^2$. We note that for any $\epsilon > 0$, we have

$$
\mu[F_n(u) > \epsilon] \leq \frac{1}{\epsilon} \int_{L^2_x} F_n(u)\, d\mu(u) = 0,
$$

where the last equality follows from (3.1) and the assumption (3.2). Letting $\epsilon \to 0$, we conclude that $\mu[F_n(u) > 0] = 0$ for all $n \in \mathbb{N}$. Equivalently, we have

$$
\mu\left[\int_D u \cdot \nabla\psi_n \neq 0\right] = 0, \quad \text{for all } n \in \mathbb{N}.
$$

Finally, we conclude that

$$
\begin{aligned}
\mu[\mathrm{div}(u) \neq 0] &= \mu\left[\bigcup_{n\in\mathbb{N}} \left\{\int_D u \cdot \nabla\psi_n \neq 0\right\}\right] \\
&\leq \sum_{n\in\mathbb{N}} \mu\left[\int_D u \cdot \nabla\psi_n \neq 0\right] \\
&= 0.
\end{aligned}
$$

Hence $\mu[\mathrm{div}(u) = 0] = 1$, i.e. μ is concentrated on divergence-free vector fields. □

Note that if $\rho, \mu \in \mathcal{P}(L^2_x)$ are probability measures, and if ρ is of the form

$$
\rho = \sum_{i=1}^M \alpha_i \delta_{u_i},
$$

where $\alpha_i > 0$, $\sum_{i=1}^M \alpha_i = 1$, and $u_i \in L^2_x$, then a transport plan from μ to ρ is necessarily of the form[14]:

$$
\pi = \sum_{i=1}^M \alpha_i \mu_i \otimes \delta_{u_i},
$$

where $\mu_i \in P(L^2_x)$, and $\sum_{i=1}^M \alpha_i\mu_i = \mu$. Therefore, given $\alpha = (\alpha_1, \dots, \alpha_M)$ as above, and $\mu \in \mathcal{P}(L^2_x)$, we denote

$$\Lambda(\alpha,\mu) := \left\{ (\mu_1,\dots,\mu_M) \,\middle|\, \mu_i \in \mathcal{P}(L^2_x), \sum_{i=1}^M \alpha_i\mu_i = \mu \right\}.$$

Note that the set $\Lambda(\alpha,\mu)$ is non-empty, since it contains $(\mu,\dots,\mu)$.

In analogy with work of Refs. 14 and 16 on entropy statistical solutions for hyperbolic systems of conservation laws, we define the following.

Definition 3.2. (Dissipative statistical solution) A statistical solution $\mu_t \in L^1_t(\mathcal{P})$ is called *dissipative*, if for every choice of coefficients $\alpha_i > 0$ with $\sum_{i=1}^M \alpha_i = 1$ and for every $(\overline{\mu}_1,\dots,\overline{\mu}_M) \in \Lambda(\alpha,\overline{\mu})$, there exists a function $t \mapsto (\mu_{1,t},\dots,\mu_{M,t}) \in \Lambda(\alpha,\mu_t)$, such that $t \mapsto \mu_{i,t}$ is weak-$*$ measurable, $\mu_{i,t}|_{t=0} = \overline{\mu}_i$, such that each $\mu_{i,t}$ satisfies

$$\int_0^T \int_{L^2_x} \int_D [u\cdot\partial_t\phi + (u\otimes u) : \nabla\phi]\, dx\, d\mu_{i,t}(u)\, dt = -\int_{L^2_x}\int_D u\cdot\phi(0,x)\, dx\, d\overline{\mu}_i(u),$$

for all $\phi \in C^\infty_c([0,T)\times D)$, $\mathrm{div}(\phi) = 0$, and all $i = 1,\dots,M$. And, in addition, we have for almost every $t \in [0,T)$:

$$\int_{L^2_x} \|u\|^2_{L^2_x}\, d\mu_{i,t}(u) \le \int_{L^2_x} \|u\|^2_{L^2_x}\, d\overline{\mu}_i(u), \quad i = 1,\dots,M.$$

3.1. *Existence and uniqueness of dissipative solutions*

As already pointed out in the introduction, the initial-value problem for the incompressible Euler equations is ill-posed for general initial data $\overline{u} \in L^2_x$, i.e. there exists an exceptional set of initial data $\mathcal{E} \subset L^2_x$ for which there either exist no suitable solutions at all, or for which there exist infinitely many suitable solutions. In practice, one may nevertheless hope that the "probability of encountering" such exceptional initial data is 0, so that the subsequent evolution would then be well defined at least for initial data encountered in practice. In this section, we provide a formal description of a suitable set of statistical initial data $\overline{\mu}$ for which this intuition holds true, and show existence and partial uniqueness of dissipative statistical solutions μ_t for initial data $\overline{\mu}$ in this class. In particular, the results in this section imply a weak–strong uniqueness result for statistical solutions. In contrast, an analogous weak–strong uniqueness result for measure-valued solutions only holds for *atomic* initial data i.e. when the initial datum is a function, but fails for non-atomic Young measure-valued initial data (see e.g. Example 1 in Ref. 13 for an explicit example in the context of conservation laws).

More precisely, we show based on purely topological arguments, that if the set of C^1-regular initial data admitting classical solutions of (1.1), over a given time-interval $[0,T)$ is dense in L^2_x, then there exists a (topologically) generic set $\mathcal{G} \subset L^2_x$, containing these regular initial data, with the following property: For any initial data $\bar{\mu} \in \mathcal{P}(L^2_x)$ which is concentrated on this generic set $\mathcal{G} \subset L^2_x$, we have *existence and uniqueness in the class of dissipative statistical solutions.* By a "generic" set $\mathcal{G}$, we denote a set whose complement $\mathcal{E} = L^2_x \backslash \mathcal{G}$ is a countable union of nowhere dense sets (implying that $\mathcal{E}$ is a meagre set in the topological sense). We say that $\bar{\mu}$ is concentrated on $\mathcal{G}$, if $\bar{\mu}(\mathcal{G}) = 1$.

The construction of such a generic $\mathcal{G}$ under the above mentioned assumption has first been carried out in Ref. 33. Let us first review the construction of $\mathcal{G}$. We let $\mathcal{C} \subset C^1(D;U)$ denote the set of initial data $\overline{v}$ admitting a classical solution $v(t)$ on $[0,T)$, with $C(\overline{v}) := \sup_{t\in[0,T)} \|\nabla v(t)\|_{L^\infty}$ finite, i.e. $C(\overline{v}) < \infty$. For $n \in \mathbb{N}$, define the open set $\mathcal{G}_n$, by

$$\mathcal{G}_n := \left\{ \overline{u} \in L^2_x \,|\, \exists\, \overline{v} \in \mathcal{C} \text{ s.t. } \|\overline{u} - \overline{v}\|_{L^2_x} < \frac{1}{n} e^{-C(\overline{v})T} \right\}. \tag{3.3}$$

Finally, we let $\mathcal{G} = \bigcap_{n\in\mathbb{N}} \mathcal{G}_n$.

Remark 3.2. If there exists a dense set of initial data $\overline{v} \in \mathcal{C}$, then $\mathcal{G}$ is generic in the topological sense (more precisely a G_δ set), being the countable intersection of the *dense* open sets $\mathcal{G}_n$. By the Baire category theorem, the set $\mathcal{G}$ is non-empty and dense in this case. In particular, this would hold true if there is no finite-time blow-up for sufficiently smooth classical solutions of the incompressible Euler equations (e.g. for $C^{1,\alpha}$ initial data $\overline{v}$ possessing a Hölder continuous derivative), which is an established fact in two space dimensions but an open question in three space dimensions.

We can now state the main theorem of the present section:

Theorem 3.1. *Define the generic set* $\mathcal{G}$ *as above* (*cf. Eq.* (3.3)). *If* $\bar{\mu} \in \mathcal{P}(L^2_x)$ *is an initial datum such that* $\bar{\mu}(\mathcal{G}) = 1$ *and there exists* $M > 0$ *such that* $\bar{\mu}(B_M(0)) = 1$, *then there exists a unique dissipative statistical solution* μ_t *of the incompressible Euler equations with initial data* $\bar{\mu}$.

The proof of Theorem 3.1 is somewhat technical and is left to Appendix F.

Remark 3.3. Even though a topologically generic set $\mathcal{G}$ is large in some sense (e.g. in the sense that a countable intersection of generic sets is again generic, and in particular non-empty), it might be very small from a different point of view. For example, there exist generic subsets of $\mathbb{R}^n$, which have Lebesgue measure zero. It is therefore *a priori* not clear whether there are any "interesting" $\bar{\mu}$, such that $\bar{\mu}(\mathcal{G}) = 1$, beyond those $\bar{\mu}$ which are concentrated on smooth initial data. Nevertheless, Theorem 3.1 implies suitable weak–strong uniqueness results as we show below.

We can derive the following corollaries from Theorem 3.1.

Corollary 3.1. (Short-time existence and uniqueness) *If* $m \geq \lfloor d/2 \rfloor + 2$, *and if there exists a* $C > 0$, *such that* $\overline{\mu} \in \mathcal{P}(L^2_x)$ *is concentrated on*

$$\{\overline{u} \in H^m_x \,|\, \|\overline{u}\|_{H^m_x} \leq C\},$$

then there exists $T^* > 0$ *(depending only on* C*) and a statistical solution* $\mu_t : [0,T^*] \to \mathcal{P}(L^2_x)$ *with initial data* $\overline{\mu}$. *Furthermore,* μ_t *is unique in the class of dissipative statistical solutions for* $t \in [0,T^*]$.

Proof. Classical short-time existence results for the Euler equations[38] show that there exists $T^* > 0$, such that for initial data $\overline{u}$ with $\|\overline{u}\|_{H^m_x} \le C$, there exists a unique solution $u(t)$ such that

$$\sup_{t\in[0,T^*]} \|u(t)\|_{H^m_x} \le C' \|\overline{u}\|_{H^m_x}.$$

Since $H^m_x \hookrightarrow C^1$, this implies that $\overline{\mu}$ is concentrated on $\overline{u} \in \mathcal{C}$. In particular, we conclude that $\overline{\mu}(\mathcal{G}) = 1$, and the result now follows from Theorem 3.1. □

Corollary 3.2. (Weak–Strong uniqueness in 2d) *Let* $d = 2$, *and let* $\alpha \in (0,1)$. *If* $\overline{\mu}$ *is concentrated on* $C^{1,\alpha}(D;U)$ *and if there exists* $M > 0$, *such that* $\overline{\mu}(B_M(0)) = 1$, *then there exists a dissipative statistical solution* μ_t *with initial data* $\overline{\mu}$. *Furthermore,* μ_t *is unique in the class of dissipative statistical solutions with initial data* $\overline{\mu}$.

Proof. Again, we observe that for any $\overline{u} \in C^{1,\alpha}$, there exists a unique solution $u(t) \in C^{1,\alpha}$. Hence, we have $\overline{u} \in \mathcal{C}$ for all such $\overline{u}$. In particular, it follows that $\overline{\mu}$ is concentrated on $\mathcal{G}$. The claim follows from Theorem 3.1. □

4. Numerical Approximation of Statistical Solutions

In this section, we will propose an algorithm for computing statistical solutions of the incompressible Euler equations (1.1). As mentioned before, this algorithm is very similar to the one proposed in Ref. 16 for computing statistical solutions of hyperbolic systems of conservation laws, which in turn was inspired by the ensemble averaging algorithms of Refs. 13 and 25 for computing measure-valued solutions. This algorithm requires a spatio-temporal discretization and a Monte Carlo sampling of the underlying probability space. We propose to use a spectral viscosity spatial discretization which is described below.

4.1. *Spectral hyper-viscosity scheme*

We write $u^\Delta(x,t) = \sum_{|k|_\infty \le N} \widehat{u}^\Delta_k(t) e^{ik\cdot x}$, where now and in the following we shall consistently denote $\Delta = 1/N$, and we denote $|k|_\infty := \max_{i=1,\dots,d} |k_i|$. We consider the following spectral viscosity approximation[50,51] of the incompressible Euler equations:

$$\begin{cases} \partial_t u^\Delta + \mathcal{P}_N(u^\Delta \cdot \nabla u^\Delta) + \nabla p^\Delta = -\epsilon_N |\nabla|^{2s}(Q_N * u^\Delta), \\ \operatorname{div}(u^\Delta) = 0, \\ u^\Delta|_{t=0} = \mathcal{P}_N \overline{u}. \end{cases} \tag{4.1}$$

Here, $\mathcal{P}_N$ is the spatial Fourier projection operator, mapping an arbitrary function $f(x,t)$ onto the first N Fourier modes: $\mathcal{P}_N f(x,t) = \sum_{|k|_\infty \le N} \widehat{f}_k(t) e^{ik\cdot x}$. Q_N is a Fourier multiplier of the form

$$Q_N(x) = \sum_{m_N < |k| \le N} \widehat{Q}_k e^{ik\cdot x} \tag{4.2}$$

and we assume $0 \le \widehat{Q}_k \le 1$.

The idea behind the SV method is that dissipation is only applied on the upper part of the spectrum, i.e. for $|k| > m_N$, thus preserving the formal spectral accuracy of the method, while at the same time enabling us to enforce a sufficient amount of energy dissipation on the small scale Fourier modes which is needed to stabilize the method. The additional hyperviscosity parameter $s \ge 1$ in (4.1) can be chosen larger to enforce more numerical dissipation on the high Fourier modes, thus allowing a larger part of the Fourier spectrum to remain free of numerical diffusion, while still ensuring stability of the resulting numerical scheme.

Note that Q_N is defined via its Fourier coefficients $\widehat{Q}_k$, which, given an additional parameter $\theta > 0$, are assumed to satisfy the following constraints:

$$\widehat{Q}_k = 0, \quad \text{for } |k| \le m_N, \quad 1 - \left(\frac{m_N}{|k|}\right)^{(2s-1)/\theta} \le \widehat{Q}_k \le 1. \tag{4.3}$$

In Ref. 51, the parameters m_N, ϵ_N, θ are chosen such that

$$m_N \sim N^\theta, \quad \epsilon_N \sim \frac{1}{N^{2s-1}}, \quad 0 \le \theta < \frac{2s-1}{2s}. \tag{4.4}$$

Multiplying the evolution equation (4.1) by u^Δ and integrating by parts, we obtain the following energy balance:

$$\|u^\Delta(t)\|^2_{L^2_x} + 2\epsilon_N \int_0^t \sum_{|k|_\infty \le N} \widehat{Q}_k |k|^{2s} |\widehat{u}^\Delta_k(\tau)|^2 \, d\tau = \|u^\Delta(t=0)\|^2_{L^2_x} \le \|\overline{u}\|^2_{L^2_x}. \tag{4.5}$$

4.2. *Monte Carlo algorithm*

Following Ref. 16, the computation of statistical solutions of (1.1) requires combining the spectral viscosity scheme (4.1) with the following Monte Carlo sampling,

Algorithm 4.1. (Monte Carlo) Given $\bar{\mu} \in \mathcal{P}(L^2_x)$, and a grid scale $\Delta = 1/N$, we determine an approximate statistical solution μ^Δ_t, as follows: For $m = m(N)$,

- Generate i.i.d. samples $\bar{u}_1, \ldots, \bar{u}_m \sim \bar{\mu}$.
- Evolve the samples, using the numerical scheme $u^\Delta_i(t) := S^\Delta_t \bar{u}_i$, where S^Δ_t denotes the solution operator, defined by the scheme (4.1).
- The approximate statistical solution μ^Δ_t is given by the so-called *empirical measure*

$$\mu^\Delta_t := \frac{1}{m} \sum_{i=1}^m \delta_{u^\Delta_i(t)}. \tag{4.6}$$

We remark that in practice, the samples $\bar{u}_i$ for $1 \leq i \leq m$ are random realizations with respect to a certain underlying probability space.

Remark 4.1. The Monte Carlo Algorithm 4.1, when restricted only to the computation of the first correlation marginal ν^1, reduces to the ensemble averaging algorithm proposed in Ref. 25 for computing measure-valued solutions of the incompressible Euler equations.

4.3. *Convergence to statistical solutions*

In this section, we will investigate the convergence of the empirical measure μ_t^Δ (4.6), generated by the Monte Carlo Algorithm 4.1, to a statistical solution of (1.1). To this end, we seek to apply the convergence Theorem 2.2 to these approximations. We start by verifying the temporal regularity of the empirical measures in the following lemma,

Lemma 4.1. *There exists $L \in \mathbb{N}$ and constants $C, C' > 0$, such that if u^Δ is obtained from the spectral hyper-viscosity method* (4.1), *with $\Delta = 1/N$ and initial data $\overline{u} \in L^2_x$, then*

$$\partial_t u^\Delta + \operatorname{div}(u^\Delta \otimes u^\Delta) + \nabla p^\Delta = E^\Delta,$$

where $\|E^\Delta\|_{H_x^{-L}} \leq C\Delta(1 + \|u^\Delta\|^2_{L^2_x})$. Furthermore, there exists a constant C', such that

$$\|u^\Delta(t) - u^\Delta(s)\|_{H_x^{-L}} \leq C'(1 + \|\overline{u}\|^2_{L^2_x})|t - s|.$$

Proof. From the definition of the spectral hyperviscosity scheme (4.1), we obtain

$$\partial_t u^\Delta + \operatorname{div}(u^\Delta \otimes u^\Delta) + \nabla p^\Delta = E^\Delta,$$

with $E^\Delta = E_1^\Delta + E_2^\Delta$, where

$$E_1^\Delta = \epsilon_N |\nabla|^{2s} u^\Delta, \quad E_2^\Delta = (I - \mathcal{P}_N)\operatorname{div}(u^\Delta \otimes u^\Delta).$$

Clearly, the first error term E_1^Δ can be estimated by

$$\|E_1^\Delta\|_{H_x^{-2s}} = \epsilon_N \||\nabla|^{2s} u^\Delta\|_{H_x^{-2s}} \leq \epsilon_N \|u^\Delta\|_{L^2_x} \leq \frac{1}{N}(1 + \|u^\Delta\|^2_{L^2_x}),$$

where we have used that $\epsilon_N \sim N^{-2s+1} \leq N^{-1}$ in the last step (assuming $s \geq 1$). On the other hand, to estimate the second term, let $\phi \in C^\infty(D; U)$ be a given vector field. Then

$$\begin{aligned}\int_D \phi \cdot E_2^\Delta \, dx &= \int_D \phi \cdot (I - \mathcal{P}_N)\operatorname{div}(u^\Delta \otimes u^\Delta)\, dx \\ &= -\int_D \nabla(I - \mathcal{P}_N)\phi : (u^\Delta \otimes u^\Delta)\, dx \\ &\leq \|(I - \mathcal{P}_N)\nabla\phi\|_{L^\infty_x} \|u^\Delta\|^2_{L^2_x}.\end{aligned}$$

Choosing $\ell > 0$ sufficiently large, we have by the Sobolev embedding theorem $H^\ell_x \hookrightarrow L^\infty_x$. Hence, we can further estimate for some absolute constant $C > 0$:

$$\begin{aligned}\|(I-\mathcal{P}_N)\nabla\phi\|_{L^\infty_x} &\le C\|(I-\mathcal{P}_N)\nabla\phi\|_{H^\ell_x} \\ &\le C\|(I-\mathcal{P}_N)\phi\|_{H^{\ell+1}_x} \\ &\le CN^{-1}\|\phi\|_{H^{\ell+2}_x}.\end{aligned}$$

Thus, we have shown that

$$\int_D \phi\cdot E_2^\Delta\,dx \le \|\phi\|_{H^{\ell+2}_x} CN^{-1}\|u^\Delta\|^2_{L^2_x},$$

for all $\phi \in C^\infty(D;U)$, which implies by duality that $\|E_2^\Delta\|_{H^{-(\ell+2)}_x} \le CN^{-1}\|u^\Delta\|^2_{L^2_x}$.

Let now $L := \max(2s, \ell+2)$. Then, combining the above two estimates, and noting that $H^{-(\ell+s)}_x, H^{-2s}_x \hookrightarrow H^{-L}_x$, we now conclude that there exists an absolute constant $C > 0$, such that

$$\|E^\Delta\|_{H^{-L}_x} \le CN^{-1}(1+\|u^\Delta\|^2_{L^2_x}).$$

Since $\Delta = 1/N$, this is the claimed estimate for E^Δ.

For the time-regularity estimate, we write

$$\partial_t u^\Delta = -\mathcal{P}_N \operatorname{div}(u^\Delta\otimes u^\Delta) - \nabla p^\Delta + E_1^\Delta.$$

By an argument analogous to the above, we then find that

$$\|\partial_t u^\Delta\|_{H^{-L}_x} = \|\mathcal{P}_N\operatorname{div}(u^\Delta\otimes u^\Delta) + \nabla p^\Delta + E_1^\Delta\|_{H^{-L}_x} \le C'(1+\|u^\Delta\|^2_{L^2_x}).$$

Recalling that also $\|u^\Delta\|_{L^2_x} \le \|\overline{u}\|_{L^2_x}$, it follows that

$$\begin{aligned}\|u^\Delta(t)-u^\Delta(s)\|_{H^{-L}_x} &= \left\|\int_s^t \partial_t u^\Delta(\tau)\,d\tau\right\|_{H^{-L}_x} \\ &\le \int_s^t \|\partial_t u^\Delta(\tau)\|_{H^{-L}_x}\,d\tau \\ &\le C'(1+\|\overline{u}\|^2_{L^2_x})|t-s|.\end{aligned}$$

This concludes our proof. □

From Lemma 4.1, it is now easy to see that if μ_t^Δ is generated by the Monte Carlo Algorithm 4.1, i.e.

$$\mu_t^\Delta = \frac{1}{M}\sum_{i=1}^M \delta_{u_i^\Delta(t)},$$

with $u_i^\Delta(t)$ computed by the spectral hyper-viscosity scheme (4.1), then the transport plan defined by

$$\pi_{s,t}^\Delta := \frac{1}{M}\sum_{i=1}^M \delta_{u_i^\Delta(s)}\otimes\delta_{u_i^\Delta(t)},$$

satisfies the properties required by the definition of time-regularity, Definition 2.2. This provides the required temporal regularity required by Theorem 2.4.

Next, we turn our attention to the spatial regularity bounds of Theorem 2.2. In particular, we need to obtain uniform estimates on the structure function (2.6). We start with the following simple observation.

Lemma 4.2. *For any* $r \geq 0$, *we have*

$$\fint_{B_r(0)} |e^{ik\cdot h} - 1|^2 \, dh \leq C \min(|k|^2 r^2, 1) \leq C|k|^2 r^2,$$

where $C = 4$.

Proof. Fix $k, h \in \mathbb{R}^d$. Let $f(\tau) := |e^{i\tau k\cdot h} - 1|$. Then

$$f(0) = 0, \quad |f'(\tau)| \leq |k||h|.$$

This implies that for $|h| \leq r$, and $\tau \in [0, 1]$:

$$|f(\tau)| \leq \int_0^t |k||h| \, ds \leq |k|r.$$

Furthermore, the upper bound $|f(t)| \leq 2$ is obvious, so that

$$|e^{ik\cdot h} - 1|^2 \leq 4 \min(|k|^2 r^2, 1).$$

Averaging over $h \in B_r(0)$, it follows that

$$\fint_{B_r(0)} |e^{ik\cdot h} - 1|^2 \, dh \leq 4 \min(|k|^2 r^2, 1). \qquad \square$$

The next result is an estimate on the structure function (2.6) at the grid scale Δ.

Lemma 4.3. *If* μ_t^Δ *is an approximate statistical solution obtained from the spectral hyper-viscosity method with* $\Delta = 1/N$, *and initial data* $\bar{\mu}$ *for which there exists* $M > 0$ *such that* $\bar{\mu}(B_M(0)) = 1$ *where* $B_M(0) = \{\|u\|_{L^2_x} < M\}$, *then*

$$S^2_\Delta(\mu_t^\Delta, T) \leq CM\Delta^{1/(2s)},$$

for some absolute constant $C > 0$. *The same estimate is also true for* $r \leq \Delta$, *i.e. we have*

$$S^2_r(\mu_t^\Delta, T) \leq CMr^{1/(2s)}, \quad \textit{for all } r \leq \Delta.$$

Proof. Our goal is to estimate the structure function $S^2_r(\mu_t^\Delta, T)$ based on the energy estimate (4.5). By construction, we have

$$\mu_t^\Delta = \frac{1}{m} \sum_{i=1}^m \delta_{u_i^\Delta(t)}$$

and u_i^Δ is obtained by the spectral hyper-viscosity method (4.1) with initial data $\overline{u}_i$. By Plancherel's theorem, we obtain

$$\int_D \fint_{B_r(0)} |u_i^\Delta(x+h) - u_i^\Delta(x)|^2 \, dh \, dx = \fint_{B_r(0)} \sum_{|k|_\infty \le N} |e^{ik\cdot h} - 1|^2 |\widehat{u}_k^\Delta|^2 \mathrm{d}h$$
$$\le C \sum_{|k|_\infty \le N} |k|^2 r^2 |\widehat{u}_{i,k}^\Delta|^2,$$

where the last estimate is shown in Lemma 4.2. Let us now fix some α with $\theta \le \alpha < 1$. We split the summation over modes $|k| \le 2N^\alpha$ and $|k| > 2N^\alpha$:

$$\int_D \fint_{B_r(0)} |u_i^\Delta(x+h) - u_i^\Delta(x)|^2 \, \mathrm{d}h \, \mathrm{d}x \le r^2 \left(\sum_{|k| \le 2N^\alpha} + \sum_{|k| > 2N^\alpha} \right) |k|^2 |\widehat{u}_{i,k}^\Delta|^2. \quad (4.7)$$

From the L^2-bound $\|u_i^\Delta(t)\|_{L_x^2} \le \|\overline{u}_i\|_{L_x^2}$, we trivially estimate

$$\sum_{|k| \le 2N^\alpha} r^2 |k|^2 |\widehat{u}_{i,k}^\Delta|^2 \le 4 r^2 N^{2\alpha} \|\overline{u}_i\|_{L_x^2}.$$

Recall that for all $|k| \ge 2N^\alpha \ge 2m_N$ we have

$$\widehat{Q}_k \ge 1 - \left(\frac{m_N}{|k|}\right)^{(2s-1)/\theta} \ge 1 - 2^{-(2s-1)/\theta} \ge 1 - 2^{-2s} = C,$$

as a consequence of (4.3) and (4.4). We thus find

$$\int_0^T \sum_{|k| > 2N^\alpha} r^2 |k|^2 |\widehat{u}_{i,k}^\Delta|^2 \, \mathrm{d}t \le C r^2 \int_0^T \sum_{|k| > 2N^\alpha} \widehat{Q}_k |k|^2 |\widehat{u}_{i,k}^\Delta|^2 \, \mathrm{d}t$$
$$\le C r^2 N^{-2\alpha(s-1)} \int_0^T \sum_{|k| > 2N^\alpha} \widehat{Q}_k |k|^{2s} |\widehat{u}_{i,k}^\Delta|^2 \, \mathrm{d}t$$
$$\le C r^2 \frac{N^{-2\alpha(s-1)} \|\overline{u}_i\|_{L_x^2}}{\epsilon_N}$$
$$= C r^2 N^{2\alpha} N^{-2\alpha s + 2s - 1} \|\overline{u}_i\|_{L_x^2},$$

where we have used the prescribed scaling $\epsilon_N \sim N^{-2s+1}$ in the last step.

Combining both estimates, and noting that for all i, $\|u_i\|_{L^2} \le M$ by assumption, we conclude that

$$\int_0^T \int_D \fint_{B_r(0)} |u_i^\Delta(x+h) - u_i^\Delta(x)|^2 \, dh \, dx \, dt \le C M r^2 N^{2\alpha} (1 + N^{-2\alpha s + 2s - 1}), \quad (4.8)$$

with an implied constant that depends on s, T, but is independent of the initial data, and N. If we now choose $r = \Delta = 1/N$, so that Δ is a length at the grid-scale, then we find

$$\int_0^T \int_D \fint_{B_\Delta(0)} |u_i^\Delta(x+h) - u_i^\Delta(x)|^2 \, dh \, dx \, dt \le C M N^{2\alpha - 2} (1 + N^{-2\alpha s + 2s - 1}).$$

We finally would like to determine a constant α, which minimizes this last expression. This is achieved when the term in brackets is of order 1, i.e. with $\alpha = \frac{2s-1}{2s}$. With this choice of α, we have $2\alpha - 2 = \frac{-1}{s}$, and

$$\int_0^T \int_D \fint_{B_\Delta(0)} |u_i^\Delta(x+h) - u_i^\Delta(x)|^2 \, dh \, dx \, dt \leq CMN^{2\alpha-2} = CMN^{\frac{-1}{s}} = CM\Delta^{1/s}.$$

Let us finally sum over $i = 1, \ldots, m$. Thus, if μ_t^Δ is an approximate statistical solution obtained from the spectral vanishing hyperviscosity method of order s, and initial data $\overline{\mu}$, with $\Delta = N^{-1}$, then it satisfies the following bound on the structure function *at the grid scale*:

$$S_\Delta^2(\mu_t^\Delta, T) \leq CM\Delta^{1/(2s)}. \tag{4.9}$$

Inspection of (4.8) shows that also

$$S_r^2(\mu_t^\Delta, T) \leq CMr^{1/(2s)},$$

if $r \leq \Delta$. □

As in Ref. 16, Sec. 4.2, we have uniform estimates on the structure function at (or below) the grid scale. Large scale features are in any case independent of the resolution Δ. However, we lack any information on the *intermediate scales*, in between the two. To close this information gap, we follow Ref. 16 and make an assumption on scaling of the structure function (2.6) at intermediate scales. The resulting theorem is the following.

Theorem 4.2. *Consider the incompressible Euler equations with initial data* $\bar{\mu} \in \mathcal{P}(L_x^2)$, *such that* $\operatorname{supp}(\bar{\mu}) \subset B_M$, *with* B_M *the ball of radius* M *in* L_x^2, *for some* $M > 0$. *Define the approximate statistical solution* μ_t^Δ *by the Monte Carlo Algorithm* 4.1. *If the approximate statistical solutions* μ_t^Δ *satisfy*:

- *Approximate scaling*: *For every* $\ell > 1$, *there exists a constant* $0 < \lambda_\ell \leq 1/(2s)$, *fixed* $C > 0$ *possibly depending on the initial data, but independent of* ℓ *and the grid size* N, *such that*

$$S_{\ell\Delta}^2(\mu_t^\Delta, T) \leq C\ell^{\lambda_\ell} S_\Delta^2(\mu_t^\Delta, T), \quad (T > 0).$$

Then the approximate statistical solutions μ_t^Δ *converge* (*up to a subsequence still denoted by* Δ), *as* $\Delta \to 0$, *to some* $\mu_t \in L_t^1(\mathcal{P})$.

Proof. By Lemma 4.3, there exists a constant $C > 0$, such that

$$S_r^2(\mu_t^\Delta, T) \leq \bar{C} r^{1/(2s)}, \tag{4.10}$$

for all $r \leq \Delta$. If $r > \Delta$, then by the assumed approximate scaling property, we write $r = \ell\Delta$, with $\ell > 1$, and obtain

$$S_r^2(\mu_t^\Delta, T) = S_{\ell\Delta}^2(\mu_t^\Delta, T)$$

$$
\begin{aligned}
&\le C\ell^{\lambda_\ell} S^2_\Delta(\mu^\Delta_t, T)\\
&\le C\bar{C}\ell^{1/2s}\Delta^{1/(2s)}\\
&= C\bar{C} r^{1/(2s)},
\end{aligned}
$$

for some constant $C\bar{C} > 0$. The convergence now follows from Theorem 2.4. □

Remark 4.2. The scaling assumption (4.10) can be interpreted as a weaker version of the scaling assumptions of Kolmogorov (see hypothesis H2, Eq. (6.3), p. 75 of Ref. 20) that was instrumental in the K41 theory for homogeneous, isotropic turbulence. In contrast to the exact scaling relation postulated by Kolmogorov, Theorem 4.2 only requires an upper bound on the structure functions; we do *not* assume (or indeed even conjecture) that the structure functions exhibit any precise scaling. The scaling assumption is fundamentally an assumption about the compactness properties (encoded in two-point correlations) of the approximate statistical solutions, stating that if we can control the smallest scales by diffusion, the large scales are expected to be reasonably well-behaved. This intuition is motivated by numerical experiments presented in Sec. 5. We also note that the inequalities in (4.10) can accommodate intermittency in the form of deviations for the standard Kolmogorov determination of the exponent 1/3 for the structure function (2.6).

Remark 4.3. (Convergence without scaling assumption) Let $\overline{\mu}$ be concentrated on a set of initial data $\mathcal{G} \subset L^2_x$, such that for any $\overline{u} \in \mathcal{G}$ there exists a strong (i.e. Lipschitz continuous) solution $u(x,t)$ for $t \in [0,T]$. Denote by $\mathscr{S}_t : \mathcal{G} \to L^2_x$ the solution operator mapping $\overline{u} \mapsto u(x,t) = \mathscr{S}_t(\overline{u})$, and let $\mathscr{S}^\Delta_t : \mathcal{G} \to L^2_x$, $\overline{u} \mapsto \mathscr{S}^\Delta_t(\overline{u})$ denote the discretized solution operator. The corresponding (exact/approximate) statistical solution is in this case given by the push-forward $\mu_t = \mathscr{S}_{t,\#}\overline{\mu}$, $\mu^\Delta_t = \mathscr{S}^\Delta_{t,\#}\overline{\mu}$. From the definition (2.4) of the metric d_T on $L^1_t(\mathcal{P})$, and the duality formula (2.3), we readily obtain the inequality

$$
d_T(\mu_t, \mu^\Delta_t) \le \int_0^T \int_{\mathcal{G}} \|\mathscr{S}_t(\overline{u}) - \mathscr{S}^\Delta_t(\overline{u})\|_{L^2_x} \, d\overline{\mu}(\overline{u})\, dt. \tag{4.11}
$$

Since $\mathscr{S}_t(\overline{u})$ is a *strong solution* for all $\overline{u} \in \mathcal{G}$, the pointwise convergence $\mathscr{S}^\Delta_t(\overline{u}) \to \mathscr{S}^\Delta_t(\overline{u})$ follows from the consistency and energy admissibility of the spectral viscosity scheme (cf. Theorem 3.2 in Ref. 25) and weak–strong uniqueness (cf. Theorem 2 in Ref. 6). Hence, the integrand on the right-hand side of (4.11) is uniformly bounded and converges to zero pointwise, as $\Delta \to 0$. By dominated convergence theorem, it follows that $\mu^\Delta_t \to \mu_t$ in $L^1_t(\mathcal{P})$. In particular, the approximate statistical solution μ^Δ_t computed by Algorithm 4.1 converges to the unique dissipative statistical solution μ_t in this case, *without any additional assumptions on the structure functions.*[a] In three dimensions, this implies the convergence approximate statistical solutions to the unique dissipative statistical solution under the assumptions of

[a]It can be shown that the convergence $\mu^\Delta_t \to \mu_t$ in turn *implies* a uniform decay of the structure functions as $\Delta \to 0$.[24]

Corollary 3.1 (short-time existence and uniqueness). In the two-dimensional case, this shows the convergence to the unique dissipative statistical solution under the assumptions of Corollary 3.2 (global existence and uniqueness for $C^{1,\alpha}$ initial data).

4.4. *Decay of energy spectrum*

In this section, we will provide an alternative criterion to ensure convergence of probability measures with respect to the metric (2.4).

This criterion is motivated from well-known experimental and theoretical concepts in the study of turbulent flows and is based on the energy spectrum $E(u;K)$ ($K \in \mathbb{N}_0$) associated to a vector field u, defined as

$$E(u;K) = \frac{1}{2} \sum_{K-1<|k|\le K} |\widehat{u}(k)|^2.$$

Note that the kinetic energy is obtained as a sum

$$\frac{1}{2}\int_D |u|^2 \, dx = \sum_{K=1}^{\infty} E(u;K).$$

Given a probability measure $\mu \in \mathcal{P}(L^2_x)$, let us similarly define:

$$E(\mu;K) = \int_{L^2_x} E(u,K)\, d\mu(u),$$

so that $E(\delta_u;K) = E(u;K)$, for $u \in L^2_x$. Finally, we denote by $E_T(\mu_t;K)$ the time-integrated energy spectrum

$$E_T(\mu_t;K) = \int_0^T E(\mu_t;K)\, dt.$$

It is an experimentally observed fact[20] that the typical energy spectrum of turbulent flows with a sufficiently strong dissipation mechanism at small scales typically takes a shape similar to the one shown in Fig. 1: Visible are three parts of the energy spectrum. The left-most part (small K) corresponds to large-scale features for the flow, the middle part (intermediate K) is referred to as the inertial range, while the right-most part (large K) may be referred to as the dissipation range. The appearance of these three parts is heuristically explained as follows. Starting from initial data (with a sufficiently fast decay of the energy spectrum) initially fixes the large-scale features of the flow. Due to the nonlinear nature of the evolution equation, these large-scale features decay to smaller scales, corresponding to energy cascading from small values of K to larger values of K. While a satisfactory mathematical treatment of the precise nature of this energy cascade remains an outstanding challenge, there is evidence by physical reasoning and as well as from numerical and real-world experiments that typically the energy spectrum resulting from this cascade process satisfies at least an upper bound of the form $E(K) \lesssim K^{-\gamma}$, for some fixed γ that is associated with the nonlinearity. In the presence of a dissipative mechanism acting on small scale features of the flow, this “free” energy cascade to larger values of K due to the nonlinearity is finally interrupted by the dissipation. Thus, energy is dissipated at dissipative scales.

From this heuristic point of view, we would expect the large-scale features to depend mostly on the initial data, while the decay of the energy spectrum at the largest values of K can be controlled in a numerical approximation scheme by a suitable choice of the numerical dissipation. On the other hand, there is no *a priori* information on the decay of the spectrum in the intermediate, *inertial* range. Hence, we make the following, rather natural, assumption,

Assumption 4.3. There exist $\beta > 0$ and constant $C > 0$ such that the computed energy spectra with Algorithm 4.1 scale as,

$$E_T(\mu_t^\Delta, K) \le CK^{-2\beta}, \quad \forall\, \Delta > 0. \tag{4.12}$$

◇

Under this assumption on the energy spectrum, we have the following convergence theorem.

Theorem 4.4. *If μ_t^Δ is obtained by the spectral viscosity method through Algorithm* 4.1, *and if the energy spectra $E_T(\mu_t^\Delta; K)$ satisfy the inertial range Assumption* 4.3 *with $\beta > 1/2$, then there exists a subsequence* (*not relabeled*) $\Delta \to 0$ *and a time-parameterized probability measure μ_t, such that $\mu_t^\Delta \to \mu_t$ in $L_t^1(\mathcal{P})$.*

Proof. From Plancherel's identity and Lemma 4.2, we have

$$\begin{aligned} S_r^2(\mu_t; T)^2 &= \int_0^T \int_{L_x^2} \fint_{B_r(0)} \int_D |u(x+h) - u(x)|^2 \, dx \, dh \, d\mu_t \, dt \\ &\lesssim \int_0^T \int_{L_x^2} \sum_k \min(|k|^2 r^2, 1) |\widehat{u}(k)|^2 \, d\mu_t \, dt \\ &\sim r^2 \sum_{K \le 1/r} K^2 E_T(\mu_t; K) + \sum_{K > 1/r} E_T(\mu_t; K). \end{aligned}$$

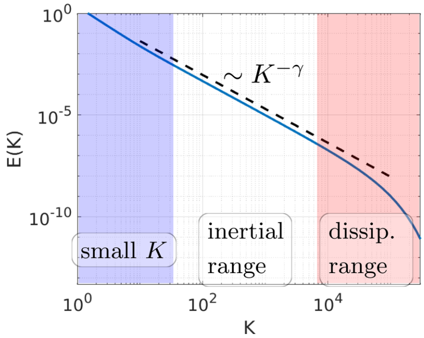
Hence, based on Assumption 4.3, we now obtain the estimate

$$\begin{aligned} S_r^2(\mu_t^\Delta; T)^2 &\lesssim r^2 \sum_{K \le 1/r} K^2 K^{-2\beta} + \sum_{K > 1/r} K^{-2\beta} \\ &\sim r^2(1 + r^{2\beta - 3}) + r^{2\beta - 1} \\ &\sim r^{\min(2, 2\beta - 1)}, \quad \text{as } r \to 0. \end{aligned}$$

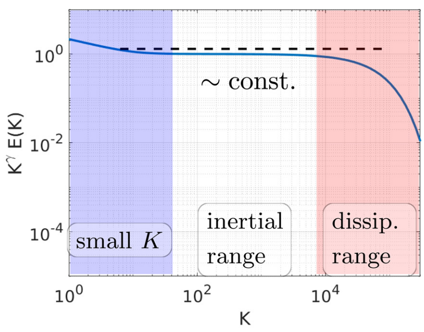
Therefore, the scaling assumption on the average energy spectrum leads to the uniform diagonal continuity:

$$E_T(\mu_t^\Delta, K) \lesssim K^{-2\beta} \Rightarrow S_r^2(\mu_t^\Delta; T) \lesssim r^{\beta - 1/2}, \quad \text{if } 1 < 2\beta < 3. \tag{4.13}$$

From Theorem 2.4, we obtain compactness of the sequence μ_t^Δ. □



(a) Energy spectrum $E(K)$



(b) Compensated E'spectrum $K^\gamma E(K)$

Fig. 1. Typical energy spectrum for turbulent flows.

Remark 4.4. As indicated in Fig. 1(b), a convenient way to check the scaling Assumption 4.3 in practice is to consider the *compensated energy spectrum*, which is defined as $K^\gamma E(K)$, where γ is the (proposed) scaling exponent in the inertial range. Proposition 4.4 says that if there exists $\gamma > 1$, such that the compensated energy spectrum $K^\gamma E_T(\mu_t^\Delta; K)$ is uniformly bounded by a constant, and independently of Δ, then $\{\mu_t^\Delta \,|\, \Delta > 0\}$ is compact in $L^1(\mathcal{P})$.

Remark 4.5. If $d = 3$ and $p = 2$, then Kolmogorov's theory states that for fully developed turbulence $S_r^2 \sim r^{1/3}$. Based on our estimate, this requires $\beta = \frac{5}{6}$. So that the (expected) energy spectrum is $E(K) \sim K^{-2\beta} \sim K^{-5/3}$. Such an assumed scaling is consistent with many real, as well as numerical, experiments reported in the literature, and is sufficient for compactness in the space of probability measures $L_t^1(\mathcal{P})$ (cf. Proposition 4.4).

4.5. *Lax–Wendroff type theorem*

We have used a compactness argument to show that under some reasonable hypotheses on the approximations, numerical solutions computed by the spectral hyper-viscosity converge to a limiting time-parameterized probability measure. In this section, we show that such a limit necessarily is a statistical solution of the incompressible Euler equations in the sense of Definition 3.1.

Theorem 4.5. (Lax–Wendroff type theorem) *Let μ_t^Δ be computed by the spectral hyper-viscosity scheme with initial data $\bar\mu$, and assume $\mu_t^\Delta \to \mu_t$ in $L_t^1(\mathcal{P})$, as $\Delta \to 0$. Then μ_t is a statistical solution of the incompressible Euler equations with initial data $\bar\mu$.*

Proof. Fix $k \in \mathbb{N}$. Let $\phi_1, \ldots, \phi_k \in C_c^\infty(D \times [0,\infty))$ be given solenoidal test functions. Set $\phi := \phi_1 \otimes \cdots \otimes \phi_k$ and denote $\nu^k = \nu^k_{x_1,\ldots,x_k,t}$. Let u^Δ be obtained from the spectral method, with initial data $\bar{u}$. Let us denote $(u,\phi) := \int_D u\cdot\phi\,dx$. Then, as a consequence of Lemma 4.1, we can write

$$\frac{d}{dt}(u^\Delta,\phi_i) = (u^\Delta,\partial_t\phi_i) + (u^\Delta\otimes u^\Delta,\nabla\phi_i) + (E^\Delta,\phi_i),$$

where there exists $L>0$ independent of Δ and the initial data $\bar{u}$, such that the error term E^Δ satisfies $\|E^\Delta\|_{H^{-L}} \le C\Delta(1+\|\bar{u}\|_{L^2_x}^2)$. Taking the product over $i=1,\dots,k$, we find

$$\begin{aligned}\frac{d}{dt}\prod_{i=1}^k(u^\Delta,\phi_i) = \sum_{i=1}^k &\left[\prod_{j\neq i}(u^\Delta,\phi_j)\right]\\ &\times\{(u^\Delta,\partial_t\phi_i) + (F(u^\Delta),\nabla\phi_i) + (E^\Delta,\phi_i)\},\end{aligned}$$

where $F(u) := u\otimes u$. Recognizing the special structure of the empirical measure μ_t^Δ (4.6) as a convex combination, denoting by $\nu^{k,\Delta} = \nu^{k,\Delta}_{x_1,\dots,x_k,t}$ the k-point correlation measure corresponding to μ_t^Δ, we obtain from the above identity that,

$$\begin{aligned}&\int_0^T\int_{D^k}\langle\nu^{k,\Delta},\xi_1\otimes\cdots\otimes\xi_k\rangle : \partial_t\phi\\ &\quad+\sum_i\langle\nu^{k,\Delta},\xi_1\otimes\cdots\otimes F(\xi_i)\otimes\cdots\otimes\xi_k\rangle : \nabla_{x_i}\phi\,dx\,dt\\ &\quad+\int_{D^k}\langle\bar{\nu}^{k,\Delta},\xi_1\otimes\cdots\otimes\xi_k\rangle : \phi(x,0)\,dx\\ &= \int_0^T\int_{L^2_x}\sum_{i=1}^k\left[\prod_{j\neq i}(u^\Delta,\phi_j)\right](E^\Delta,\phi_i)\,d\mu_t^\Delta\,dt.\end{aligned}\tag{4.14}$$

The right-hand side can be bounded by

$$\int_0^T\int_{L^2_x}\|u^\Delta\|_{L^2_x}^{k-1}\sum_{i=1}^k\prod_{j\neq i}\|\phi_j\|_{L^2_x}\|E^\Delta\|_{H_x^{-L}}\|\phi_i\|_{H_x^L}\,d\mu_t^\Delta\,dt,$$

which, by Lemma 4.1 is further bounded by

$$\le C(\phi,k)\Delta\int_0^T\int_{L^2_x}(1+\|u^\Delta\|_{L^2_x}^2)^k\,d\mu_t^\Delta(u)\,dt.$$

Note that if $\bar{\mu}$ is supported on $\overline{B_M(0)}\subset L^2_x$, then it follows that μ_t^Δ is supported on $\overline{B_M(0)}$, as well. This is a consequence of the *a priori* L^2-bound (4.5). Hence the error term in Eq. (4.14) is in this case bounded by $C\Delta$, where $C=C(\phi,k,M,T)$ is a constant independent of Δ.

Let us also note that the terms on the left-hand side of (4.14) converge strongly in $L^1_{t,x}$ as $\Delta\to 0$. Indeed, it is not difficult to see that all terms on the left-hand side, e.g.

$$g(t,x,\xi) := (\xi_1\otimes\cdots\otimes F(\xi_i)\otimes\cdots\otimes\xi_k) : \nabla_{x_i}\phi(x,t),$$

are admissible observables in the sense of (2.12). For such observables, the $L^1_{t,x}$-convergence of

$$\langle \nu^{k,\Delta}_{t,x}, g(t,x,\xi)\rangle \to \langle \nu^k_{t,x}, g(t,x,\xi)\rangle, \quad \text{as } \Delta \to 0,$$

has been established in Theorem 2.4. The same holds true for the other two terms on the left-hand side.

Passing to the limit $\mu^\Delta_t \to \mu_t$, it thus follows that

$$\begin{aligned}\int_0^T \int_{D^k} &\langle \nu^k_{t,x}, \xi_1 \otimes \cdots \otimes \xi_k\rangle : \partial_t \phi \\ &+ \sum_i \langle \nu^k_{t,x}, \xi_1 \otimes \cdots \otimes F(\xi_i) \otimes \cdots \otimes \xi_k\rangle : \nabla_{x_i}\phi \, dx\, dt \\ &+ \int_{D^k} \langle \bar\nu^k_x, \xi_1 \otimes \cdots \otimes \xi_k\rangle : \phi(x,0)\, dx = 0.\end{aligned}$$

The fact that μ_t is concentrated on incompressible vector fields follows immediately from the corresponding property of the approximations μ^Δ_t (cf. Lemma 3.1). Furthermore, from Proposition 2.2, it also follows that the limit μ_t is time-regular. This finishes the proof that μ_t is a statistical solution of the incompressible Euler equations with initial data $\bar\mu$. □

Remark 4.6. It is straightforward to show that if μ^Δ_t are generated from the spectral hyper-viscosity scheme (4.1), and if they satisfy the assumptions of Theorem 4.5, the limit μ_t is in fact a dissipative statistical solution in the sense of Definition 3.2.

5. Numerical Experiments

In this section, we will present a suite of numerical experiments to demonstrate the effectiveness of the Monte Carlo Algorithm 4.1 in computing statistical solutions of the incompressible Euler equations.

5.1. *Implementation*

For our numerical experiments, we use the implementation of the spectral hyper-viscosity scheme (4.1) provided by the SPHINX code, which was first presented in Ref. 27. In SPHINX, the nonlinear advection term is implemented using the $O(N^2 \log N)$-costly fast Fourier-transform. Aliasing is avoided via the use of a padded grid, as e.g. described in Refs. 27 and 26. The spectral scheme is implemented based on the primitive variable formulation.

For the numerical experiments reported below, we use a spectral viscosity operator of order $s = 1$ (cf. Eq. (4.1)), with $\epsilon_N = \epsilon/N$, $\epsilon = 1/20$ unless otherwise stated.

The Fourier multiplier Q_N is chosen with Fourier coefficients

$$\widehat{Q}_k = \begin{cases} 1 - N/|k|^2 & |k| \geq \sqrt{N}, \\ 0, & \text{otherwise.} \end{cases}$$

corresponding to $m_N = \sqrt{N}$.

For each sample, SPHINX solves the following system of ODEs for the Fourier coefficients $\hat{u}_k(t)$, $|k|_\infty \leq N$ of $u(x,t) = \sum_{|k|_\infty \leq N} \widehat{u}_k(t) e^{ikx}$:

$$\begin{cases} \dfrac{d}{dt}\widehat{u}_k = -ik \cdot \left(I - \dfrac{k \otimes k}{|k|^2} \right) \cdot (\widehat{u \otimes u})_k - \dfrac{\epsilon}{N} \max(|k|^2 - N, 0)\, \widehat{u}_k, \\ \widehat{u}_k(0) = \widehat{\overline{u}}_k, \end{cases} \tag{5.1}$$

for $|k|_\infty \leq N$, $k \neq 0$. Here, $\widehat{\overline{u}}_k$ denote the Fourier coefficients of the initial data. The multiplication by the matrix $(I - (k \otimes k)/|k|^2)$ implements the Leray projection onto divergence-free vector fields. The dot product with $ik \cdot (\ldots)$ corresponds to taking the divergence in Fourier space. We note that the zeroth Fourier component of u is constant in time, reflecting conservation of momentum. In SPHINX, this component is set equal to $\widehat{u}_0 \equiv 0$. The above system of ODEs (5.1) is integrated in time using an adaptive explicit third-order Runge–Kutta method.

Although the theory of Sec. 4 is valid for both two and three space dimensions and the SPHINX code is available for both cases, we restrict our focus to two space dimensions in this section, on account of affordable computational costs.

5.2. *Flat vortex sheet*

Vortex sheets occur in many models in physics and are an important test bed for numerical experiments for the Euler equations, Ref. 25 and references therein. We first consider a randomly perturbed version of the *flat vortex sheet* that corresponds to the following initial data also considered in Ref. 25,

5.2.1. *Initial data*

Given a smoothing parameter $\rho > 0$, and a parameter $\delta \geq 0$ (measuring the size of the random perturbation of the interface), this vortex sheet initial data is of the form

$$\overline{u}^{\rho,\delta}(x) = \mathbb{P}(U^\rho(x_1, x_2 + \sigma_\delta(x_1))), \tag{5.2}$$

where $\mathbb{P}$ denotes the Leray projection, $U^\rho(x) = (U_1^\rho(x), U_2^\rho(x))$ is the following smoothened flat vortex sheet initial data:

$$U^\rho(x) := \begin{cases} \tanh\left(\dfrac{x_2 - 1/4}{\rho}\right), & (x_2 \leq 1/2), \\ \tanh\left(\dfrac{3/4 - x_2}{\rho}\right), & (x_2 > 1/2), \end{cases} \qquad U_2^\rho(x) = 0.$$

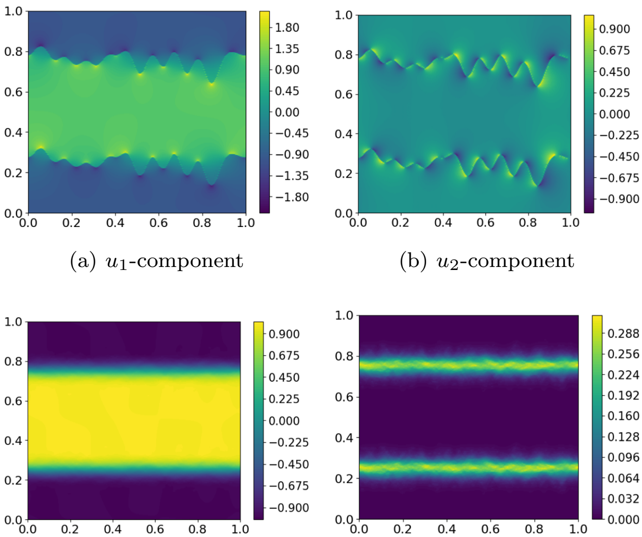
and $\sigma_\delta(x)$ is a random function, which for a given (random) choice of parameters $\alpha_1, \dots, \alpha_q \in (0, \delta)$, $\beta_1, \dots, \beta_q \in [0, 2\pi)$, is defined by

$$\sigma_\delta(x_1) = \sum_{k=1}^{q} \alpha_k \sin(2\pi x_1 - \beta_k). \tag{5.3}$$

We will also consider the discontinuous case of initial data that are obtained in the limit $\rho \to 0$ resulting in

$$U_1^0(x) := \begin{cases} +1, & (1/4 < x_2 \le 3/4), \\ -1, & (\text{otherwise}), \end{cases} \qquad U_2^0(x) = 0.$$

For our simulations, we fix $q = 10$ modes for the perturbations. The coefficients α_k are drawn independently, uniformly in $(0, 1)$, and then multiplied by δ. The coefficients β_k are i.i.d., with a uniform distribution on $[0, 2\pi)$. The initial data for the statistical solution $\bar{\mu}_\rho^\delta \in \mathcal{P}(L_x^2)$ is defined as the law of these random perturbations. It depends on the two parameters $\rho \ge 0$, $\delta \ge 0$. While ρ controls the smoothness of the initial data, δ measures the amplitude of the perturbation. We fix $\delta = 0.025$ in the following and consider different values of ρ. Note that the choice $\rho = 0$ corresponds to an initial measure supported on discontinuous flows with a very sharp transition (see Figs. 2(a) and 2(b) for realizations (samples) of this initial data). In Figs. 2(c) and 2(d), we present the initial mean and variance that correspond to the random variations of the initial interface location.



(a) u_1-component (b) u_2-component

(c) mean: u_1-comp. (d) variance: u_1-comp.

Fig. 2. Initial data for the perturbed discontinuous flat vortex sheet ($\rho = 0$), samples for $u_{1,2}$, and mean and variance of u_1.

Clearly when $\rho > 0$, the corresponding initial data for every sample is smooth. Consequently, smooth solutions of (1.1) are well-posed and the spectral viscosity method converges to this solution as $N \to \infty$.[2] However for $\rho = 0$, which corresponds to the case of a discontinuous vortex sheet, there are no well-posedness results even for weak solutions, as the vorticity corresponding to the initial datum (for each sample) is a sign changing measure and does not belong to the *Delort Class*. In Ref. 25, the authors had presented multiple numerical experiments to illustrate the approximate solutions, computed with a spectral viscosity method, may not converge (or converge too slowly to be of practical interest) for individual samples (see Figs. 5 and 6 of Ref. 25). Hence, it would be interesting to study if approximate statistical solutions, generated by Algorithm 4.1 converge in this case.

5.2.2. *Structure functions and Compensated Energy spectra*

The convergence Theorem 4.2, based on the compactness Theorem 2.2, provides us with verifiable criteria to check convergence of Algorithm 4.1. In particular, we need to check certain decay conditions on the structure function (2.6) for small correlation lengths. To this end, we consider the following instantaneous version of the structure function (2.6),

$$S^{2,\Delta}_{r,t}(\mu_t) := \left(\int_{L^2_x} \int_D \fint_{B_r(0)} |u(x+h) - u(x)|^2 \, dh \, dx \, d\mu^\Delta_t(u) \right)^{1/2}, \tag{5.4}$$

Note that the above is a formal definition and it can be made rigorous in terms of the time-dependent correlation measures. It is much simpler to compute the instantaneous quantity (5.4) than the time-averaged version (2.6).

Our objective is to check whether the structure function (2.6), or rather its instantaneous version (5.4), decays (uniformly in resolution Δ) as $r \to 0$. Such a decay would automatically imply convergence of the approximations to a statistical solutions by Theorems 2.2 and 4.5.

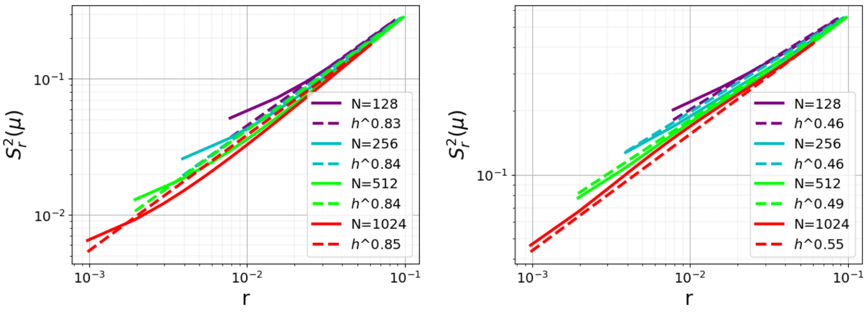
Clearly if $\rho > 0$ in (5.2), the spectral viscosity method converges to the unique classical solution as $\Delta \to 0$. Moreover, a straightforward calculation shows that the structure function (5.4) should scale as

$$S^{2,\Delta}_{r,t}(\mu_t) \approx r, \quad \forall\, \Delta, t. \tag{5.5}$$

This is indeed verified from Fig. 3(a) where we plot the structure function (5.4) at $t = 0.4$ and $\rho = 0.1$ for different values of the mesh parameter. We see from this figure that $S^{2,\Delta}_{r,0.4}(\mu_t) \approx r^{0.9}$, at fine resolutions, which is very close to the expected value of 1 for the scaling exponent of the structure function.

On the other hand, for $\rho = 0$, corresponding to the discontinuous flat vortex sheet, the lack of smoothness inhibits us from inferring a particular form of decay of (2.6) (or (5.4)) *a priori*.

At least initially, calculations in Ref. 24 imply that $S^{2,\Delta}_{r,0}(\bar{\mu}) \sim r^{\frac{1}{2}}$, for the discontinuous flat vortex sheet. Surprisingly, we find from Fig. 3(b) that at fine



(a) $\rho = 0.1$ (b) $\rho = 0.0$

Fig. 3. Instantaneous structure function (5.4) versus correlation length r for different resolutions ($N \sim \Delta^{-1}$) for different values of smoothness parameter ρ, at $t = 0.4$.

resolutions, $S^{2,\Delta}_{r,0.4}(\mu_t) \approx r^{0.52}$, which agrees with the decay of the structure function of the initial data. Although we do not present the results there, we observe that the structure function (5.4) scales as r^{θ_t}, with $\theta_t \geq 0.5$ for all t. This implies an uniform decay of the structure function (2.6) and convergence of the approximations to a statistical solution of the Euler equations (1.1), even for this case of discontinuous vortex sheet data. Note that the computed structure functions (5.4) in Fig. 3 clearly satisfy the approximate scaling hypothesis (4.10) and thus imply convergence through Theorem 4.2.

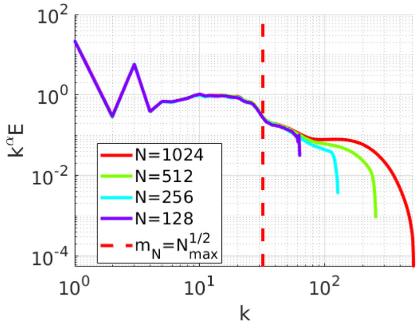
An alternative criterion for convergence of statistical solutions is provided by the energy spectrum decay in the inertial range (4.12). To check whether this criterion is satisfied, we follow Theorem 4.4 and compute the following instantaneous *compensated energy spectrum*

$$\mathcal{C}^{\Delta}_{\gamma,t}(\mu_t; K) := K^{\gamma} E(\mu_t, K). \tag{5.6}$$

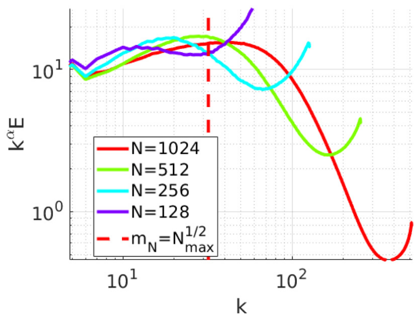
Following the arguments in the proof of Theorem 4.4, we can relate the decay of the instantaneous energy spectrum to the corresponding decay of the structure function (5.4) by a direct analogue of (4.13).

For $\rho = 0.1$ in (5.2), we plot the compensated energy spectrum $\mathcal{C}^{\Delta}_{3,0.4}(\mu_t; K)$ for all K and at time $t = 0.4$, with compensating factor $\gamma = 3$ in Fig. 4(a). Note that this choice of γ is consistent with a decay exponent of 1 for the structure function in (4.13), i.e. $S^2_r(\mu^{\Delta}_t; T) \lesssim r$. We observe from this figure that as expected for this case, the compensated energy spectrum is clearly bounded and in fact, decays faster than the expected rate for the entire range of wave numbers.

On the other hand, we plot the compensated energy spectrum $\mathcal{C}^{\Delta}_{2,0.4}(\mu_t; K)$ (5.6) for the discontinuous flat vortex sheet case, i.e. $\rho = 0$ in (5.2), in Fig. 4(b). In this case, we expect from the structure function computations (see Fig. 3(b)) that the instantaneous structure function decays with an exponent of ≈ 0.5. From (4.13), we see that this corresponds to the choice of $\gamma = 2$ as the exponent of compensation in (5.6). Moreover, in Fig. 4(b), we also plot the line corresponding to wave number



(a) $\rho = 0.1, \gamma = 3$



(b) $\rho = 0.0, \gamma = 2$

Fig. 4. The instantaneous compensated energy spectrum $\mathcal{C}^{\Delta}_{\gamma,t}(\mu_t; K)$ (5.6) for the flat vortex sheet, at time $t = 0.4$. Note different values of γ for the smooth and discontinuous vortex sheets.

$m_N \approx \sqrt{N}$, which for the spectral viscosity method (4.1) represents the wave number after which the spectral viscosity is activated and hence, demarcates the separation between inertial and dissipation ranges. We observe from Fig. 4(b) that the compensated energy spectrum is clearly uniformly bounded (in terms of the resolution Δ) for the whole inertial range and for all resolutions Δ barring the coarsest resolution, and decays fast in the dissipation range, although there is a slight kink upwards at the very end of the dissipation range, almost at the grid scale. This might be attributed to numerical errors, which are dominant at this range. Translating these results to the energy spectrum, we see that the spectrum decays as K^{-2} in the inertial range uniformly with respect to resolution. Hence, according to Theorems 4.4 and 4.5, the sequence of approximations will converge to a statistical solution of (1.1).

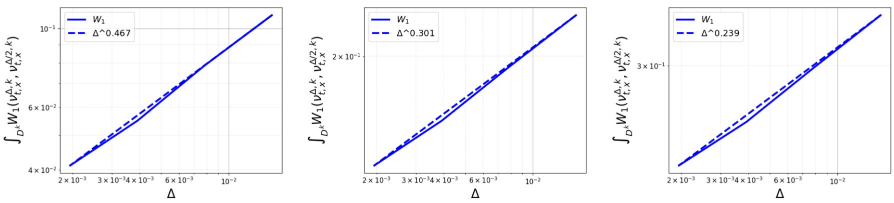
5.2.3. *Convergence in Wasserstein Metrics*

Given the computational results on the structure function and the compensated energy spectra, results in Sec. 4 clearly imply convergence of the approximations μ_t^{Δ}, generated by the Monte Carlo Algorithm 4.1 to a statistical solution of the incompressible Euler equations. Moreover from the discussion in Sec. 2, we should observe with respect to the following *Cauchy rates*:

$$d_T(\mu_t^{\Delta}, \mu_t^{\Delta/2}) = \int_0^T W_1(\mu_t^{\Delta}, \mu_t^{\Delta/2})\, dt. \tag{5.7}$$

Unfortunately, the calculation of the Wasserstein distance between probability measured defined on high-dimensional (or indeed ∞-dimensional) spaces is a highly non-trivial issue, which we cannot tackle with present computational resources.

On the other hand, one can compute finite-dimensional *marginals* of (5.7) by utilizing the complete characterization of $L^1_t(\mathcal{P})$ in terms of *correlation measures*



(a) $k = 1$ (b) $k = 2$ (c) $k = 3$

Fig. 5. The Wasserstein distances between correlation marginals $\int_{D^k} W_1(\nu^{\Delta,k}_{t,x}, \nu^{\Delta/2,k}_{t,x})\, dx$ for $k = 1, 2, 3$, at time $t = 0.4$ with respect to resolution.

as given in Theorem 2.3. Following Ref. 16 (Theorem 5.7), one can prove that

$$\int_{D^k} W_1(\nu^{\Delta,k}_{t,x}, \nu^{\Delta/2,k}_{t,x})\, dx \le C_k W_1(\mu^{\Delta}_t, \mu^{\Delta/2}_t), \quad \text{a.e. } t \tag{5.8}$$

Here, $k \ge 1$ and $\nu^{\Delta,k}_{t,x}$ is the kth correlation marginal corresponding to the approximate statistical solution μ^{Δ}_t. Note that we consider instantaneous versions of the Wasserstein metric (5.7) for reasons of computational convenience.

We remark that computing the Wasserstein distances $W_1(\nu^{\Delta,k}_{t,x}, \nu^{\Delta/2,k}_{t,x})$ for small k is much more tractable. We have computed these Wasserstein distances using the algorithm of Ref. 5 (as implemented in Ref. 18) and the corresponding results for $k = 1, 2, 3$, at time $t = 0.4$ for the discontinuous flat vortex sheet, i.e. $\rho = 0$ in (5.2) are presented in Fig. 5. As seen from this figure, we observe a clear convergence of these Wasserstein distances (in the Cauchy sense as in (5.8)) for the one-point, two-point and three-point correlation measures, albeit at a slow rate for the second and third correlation marginals. This, together with the results on the structure function and compensated energy spectra, provides considerable evidence that the approximate statistical solutions, generated by Algorithm 4.1, converge to a statistical solution of (1.1). Moreover, given Theorem 2.4, results shown in Fig. 5 establish convergence with respect to any admissible observable in the sense of (2.12), corresponding of one-point, two-point and three-point statistical quantities of interest. These include mean, variance, structure functions, energy spectra as well as three-point correlation functions.

5.3. *Sinusoidal vortex sheet*

In this section, we will consider a random perturbation of the so-called sinusoidal vortex sheet, i.e. the initial vorticity is concentrated on a sine curve. This test case was extensively studied in a recent paper[26] in the context of the numerical approximation of weak solutions (in Delort class) of the two-dimensional incompressible Euler equations. Whereas Ref. 26 considered the deterministic problem with fixed initial data, we will here follow a statistical approach, considering an initial measure $\overline{\mu}$ supported on small random perturbations of the sinusoidal vortex sheet. As discussed in Ref. 26, due to inherent Kelvin–Helmholtz instabilities the computed numerical approximations for sinusoidal vortex sheet initial data experience vortex sheet roll-up at ever smaller length-scales at increasing resolution $\Delta \to 0$ (and at low diffusivity). These small-scale Kelvin–Helmholtz instabilities slow down, and at even smaller values of Δ ultimately prevent the strong convergence of the numerical approximants to a limiting solution. In this section, we will compare the convergence properties of the deterministic problem with the corresponding perturbed statistical approach. In contrast to the deterministic problem, the quantities of interest in the statistical setting, such as the mean, variance (as well as higher-order correlations), appear to retain some smoothness even after the complex vortex sheet roll-up. This makes them amenable to numerical approximation, even though the deterministic evolution cannot be stably resolved.

5.3.1. *Initial Data*

We fix a sinusoidally perturbed vortex sheet, where the initial vorticity is a Borel measure of the form

$$\omega_0 = \delta(x - \Gamma) - \int_{\mathbb{T}^2} d\Gamma,$$

such that $\int_{\mathbb{T}^2} \omega_0 \, dx = 0$, and up to a constant, ω_0 is uniformly distributed along a curve Γ, which is defined as the graph:

$$\Gamma = \{(x, y) \,|\, y = d \sin(2\pi x), x \in [0, 1]\}.$$

We chose $d = 0.2$ for our simulations.

The numerical initial data is obtained from the mollification of this initial data with a parameter $\rho > 0$. As a mollifier, we consider the third-order B-spline

$$\psi(r) := \frac{80}{7\pi}[(r+1)_+^3 - 4(r+1/2)_+^3 + 6r_+^3 - 4(r-1/2)_+^3 + (r-1)_+^3].$$

Next, we define $\psi_\rho(x) = \rho^{-2}\psi(|x|/\rho)$. The numerical approximation of the perturbed vortex sheet is now defined by setting

$$\omega_0^\rho(x) := \int_\Gamma \psi_\rho(x - y)\omega_0(y) \, dy,$$

where ρ determines the thickness (smoothness) of the approximate vortex sheet. The convolution at $x = (x_1, x_2) \in \mathbb{T}^2$ is evaluated via numerical quadrature:

$$(\omega_0 * \psi_\rho)(x) \approx \frac{\rho}{Q} \sum_i \psi_\rho(x - (\xi_i, g(\xi_i)))\sqrt{1 + |g'(\xi_i)|^2},$$

with $\xi_i = x_1 + i\rho/Q$ equidistant quadrature points in x_1, and $g(\xi)$ the function whose graph is Γ, i.e. $g(\xi) = d \sin(2\pi\xi)$. We choose $Q = 400$ quadrature points. We denote by $U^\rho(x_1, x_2)$ the velocity field such that $\mathrm{div}(U^\rho) = 0$ and $\mathrm{curl}(U^\rho) = \omega_0^\rho$.

Similar to the case of the flat vortex sheet, we carry out random perturbations of the sinusoidal vortex sheet as follows:

$$\overline{u}^{\rho,\delta}(x) := \mathbb{P}(U^\rho(x_1, x_2 + \sigma_\delta(x_1)).$$

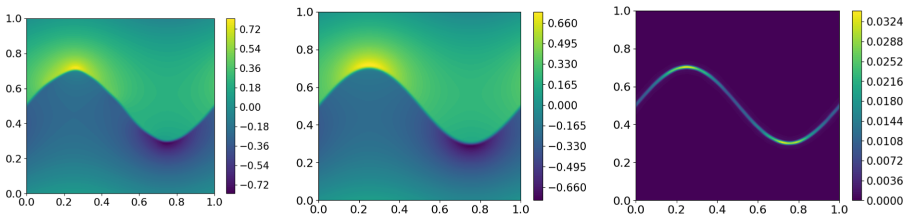
Here, $\mathbb{P} : L^2_x \to L^2_x$ denotes the Leray projection onto divergence-free vector fields, and we again fix a random function $\sigma_\delta(x)$,

$$\sigma_\delta(x_1) = \sum_{k=1}^{q} \alpha_k \sin(2\pi x_1 - \beta_k),$$

depending on a parameter $\delta \geq 0$ and a choice of (random) coefficients $\alpha_1, \dots, \alpha_q \in (0, \delta)$, $\beta_1, \dots, \beta_q \in [0, 2\pi)$. For our simulations, we fix $q = 10$ modes for the perturbations. In practice, the coefficients α_k are first drawn independently, uniformly in $(0, 1)$, and then multiplied by δ. The coefficients β_k are i.i.d., with a uniform distribution on $[0, 2\pi)$. The initial data for the statistical solution $\bar{\mu}^\delta_\rho \in \mathcal{P}(L^2_x)$ is defined as the law of these random perturbations. It depends on the two parameters $\rho \geq 0$, $\delta \geq 0$. While ρ controls the smoothness of the initial data, δ measures the amplitude of the perturbation. We fix $\delta = 0.003125$ in the following and vary ρ as a function of the grid size N. To approximate vortex sheet initial data, we must scale $\rho = \rho(N)$ with N, such that $\rho \to 0$ as $N \to \infty$. We use $\rho = 5/N$ for our simulations. The additional diffusion parameter ϵ of the spectral viscosity scheme is set to $\epsilon = 0.01$. With this choice of parameters, we will drop the sub- and superscripts and denote the initial data at a given resolution simply by $\overline{\mu} \in \mathcal{P}(L^2_x)$.

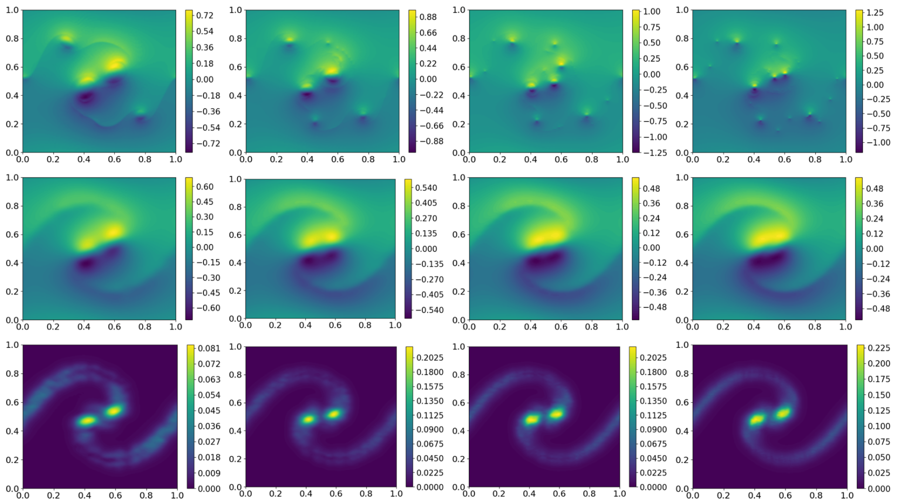
5.3.2. *Computation of individual samples*

For any single realization of the random perturbation $\sigma_\delta(x)$, the resulting vorticity of the initial datum (sample) is a positive measure, concentrated on a sine curve (see Fig. 6(a) for horizontal component of velocity u_1). Hence, any single sample of the initial data in the Delort class. Therefore, by the results of Ref. 26, the approximate solutions generated by the spectral viscosity method (4.1) will converge, on increasing resolution, to a weak solution of (1.1). However, as noted in Ref. 26, this convergence can be very slow as the flow breaks down into smaller and smaller vortices. In fact, this phenomenon is also seen from Fig. 7 (top row), where we plot the horizontal component of velocity u_1 at time $t = 1.2$ and different resolutions. At this time, the initial vortex sheet has rolled over and broken down into a succession of small vortices, whose location and amplitude are different for different resolutions. This very slow convergence is also displayed in Fig. 8(a), where we plot the Cauchy



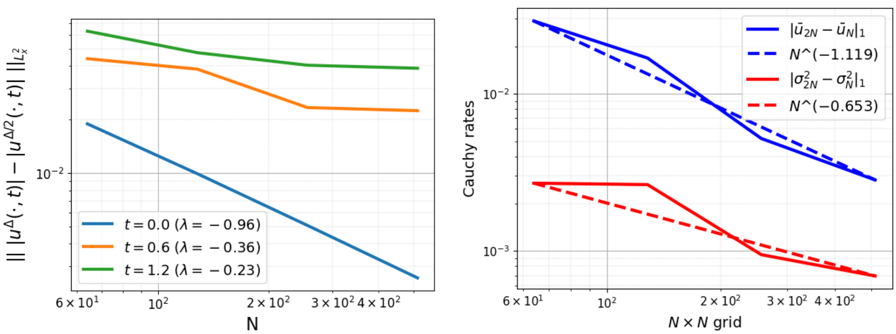
(a) Sample (b) Mean (c) Variance

Fig. 6. Initial conditions for the horizontal velocity u_1 for the sinusoidal vortex sheet.



(a) $N = 128$ (b) $N = 256$ (c) $N = 512$ (d) $N = 1024$

Fig. 7. Results at time $T = 1.2$ for the horizontal velocity u_1 of the sinusoidal vortex sheet, at different resolutions. Top row: Sample; middle row: Mean; bottom row: Variance.



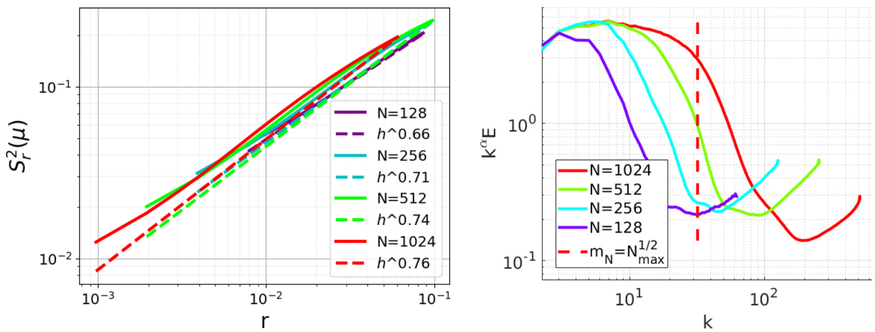
(a) Sample (b) Mean and Variance

Fig. 8. Cauchy rates for the norm of the velocity field ($\sqrt{u_1^2 + u_2^2}$) for the sinusoidal vortex sheet. Slope λ is determined by a best fit. Left: Sample convergence rates at three different times $t = 0, 0.6, 1.2$. Right: Convergence of mean and variance at $T = 1.2$.

rates $\|u^{\Delta}(t) - u^{\Delta/2}(t)\|_{L_x^2}$, with u^{Δ} denoting the approximate solution computed with the spectral viscosity method (4.1), for three different times $t = 0, 0.6, 1.2$. As seen from this figure, the rate of convergence decreases very rapidly and at time $t = 1.2$, it appears as if there is no convergence on mesh refinement.

5.3.3. *Structure functions and Compensated energy spectra*

Given this apparent non-convergence of individual samples, it is pertinent to investigate if computing the statistics will be more convergent. To this end, we consider



(a) Instant. structure function (5.4)

(b) Compensated energy spectrum (5.6) with $\gamma = 2.2$

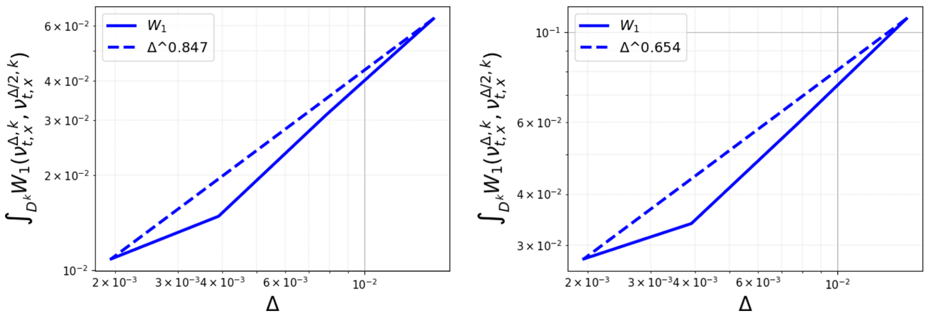
Fig. 9. The structure function and compensated energy spectrum for the sinusoidal vortex sheet at time $T = 1.2$.

the initial data to be the initial probability measure $\bar{\mu}$. The mean and variance (of the horizontal component u_1) are plotted in Figs. 6(b) and 6(c). From this figure, we observe that the initial probability measure is concentrated on very small perturbations of the underlying sinusoidal vertex sheet, as reflected in the initial variance.

In order to investigate the convergence of approximations to the statistical solution, generated by Algorithm 4.1, we follow the template of the previous numerical experiment and compute the (instantaneous) structure function (5.4) and the compensated energy spectrum (5.6) in Fig. 9. From this figure, we observe that the structure function at time $t = 1.2$ scales with an exponent of ≈ 0.7 at the finest resolutions. From (4.13), this implies roughly a $\gamma = 2.4$ in the scaling of the energy spectrum (5.6). A better fit to the scaling of the energy spectrum is found with $\gamma = 2.2$. We plot the compensated energy spectrum with the latter value of γ in Fig. 9(b). From this figure, we see that for the inertial range, the energy spectrum clearly decays (faster than) a rate of 2.2. Thus, the assumptions of Theorems 2.2 and 4.4 are satisfied and the approximations will converge to a statistical solution of (1.1).

5.3.4. *Convergence of observables and Wasserstein Distances*

Given the results on the computed structure functions and energy spectra, the approximations will converge. But is this convergence at a better rate than that of single samples? To investigate this issue, we consider two different sets of computations. First, we compute the mean and the variance of the velocity field at different resolutions and plot them (for the horizontal velocity at time $t = 1.2$) in Fig. 7(middle and bottom rows). Clearly, the one-point statistics appear much more convergent than the single sample results. The mean flow consists of a coherent set of large vortices, which is in stark contrast to the large number of vortices formed



(a) $k = 1$ (b) $k = 2$

Fig. 10. The metrics $\int_{D^k} W_1(\nu^{\Delta,k}_{t,x}, \nu^{\Delta/2,k}_{t,x})dx$ for the one- and two-point correlation marginals for the sinusoidal vortex sheet at time $t = 1.2$.

in the single sample simulations. Moreover, we also plot Cauchy rates for the mean and the variance, corresponding to the norm $\sqrt{u_1^2 + u_2^2}$ at time $t = 1.2$, and different resolutions in Fig. 8(b). Again, we observe that these one-point statistics converge at a significantly faster rate than the single sample. These results indicate that one can expect significantly better convergence of approximations for statistics than for individual realizations of fluid flows, even if the initial probability measure is a small perturbation of the underlying deterministic data and further reinforces the results of Refs. 13, 25 and 16 in this direction.

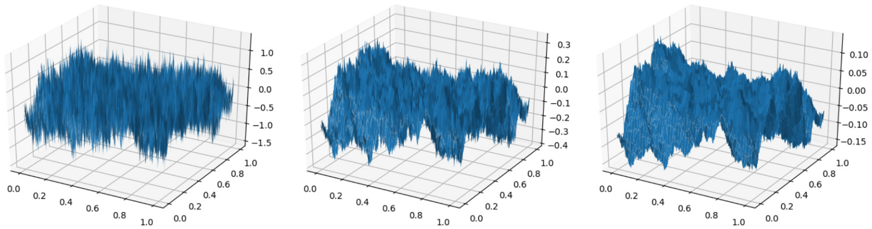
Finally, we plot the Wasserstein distances (5.8) for $k = 1, 2$, corresponding to the one- and two-point correlation marginals, at time $t = 1.2$, in Fig. 10. The results clearly show convergence in these metrics at a significantly faster rate than for individual samples and indicate possible convergence in the metric (2.4) on probability measures on L^2.

5.4. *Fractional Brownian motion*

The study of the evolution of initial ensembles corresponding to (fractional) Brownian motion stems from Refs. 46 and 48, where the authors model interesting aspects of Burgers turbulence by evolving Brownian motion initial data for the (scalar) Burgers' equation, see Ref. 15 for a more recent numerical study. Similarly in Ref. 16, the authors consider the compressible Euler equations with (fractional) Brownian motion initial data. Following these papers, we will consider the two-dimensional Euler equations (1.1) with initial data corresponding to fractional Brownian motion, i.e. the following initial data:

$$u_0^{x,H}(\omega; x) := B_1^H(\omega; x), \quad w_0^{y,H}(\omega; x) := B_2^H(\omega; x). \quad \text{for } \omega \in \Omega, \ \ x \in D, \tag{5.9}$$

where B_1^H and B_2^H are two independent two-dimensional fractional Brownian motions with the Hurst index $H \in (0, 1)$. Standard Brownian motion corresponds to a Hurst index of $H = 1/2$. The initial probability measure $\bar{\mu}$ is the law of the above random field.



(a) $H = 0.15$ (b) $H = 0.5$ (c) $H = 0.75$

Fig. 11. A single sample of initial horizontal velocity u_1 for the fractional Brownian motion initial data (5.9) for three different Hurst indices.

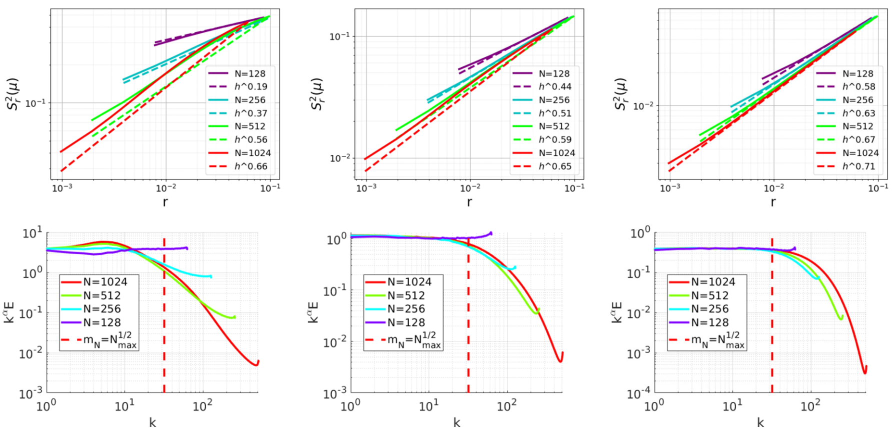
To generate fractional Brownian motion, we use the random midpoint displacement method originally introduced by Lévy[29] for Brownian motion, and later adapted for fractional Brownian motion, see Sec. 6.6.1 in Ref. 16.

Considering fractional Brownian motion initial data (5.9) is a significant deviation from the vortex sheet initial data in the following respects:

- For the vortex sheet initial data, the initial measure $\bar{\mu} \in \mathcal{P}(L^2_x)$ was concentrated on a 20-dimensional subset of L^2_x (corresponding to the choice of 20 free parameters α_k, β_k). On the other hand, in the limit of infinite resolution ($\Delta \to 0$), the fractional Brownian motion initial data corresponds to a measure concentrated on an infinite dimensional subset of L^2_x.
- For any $0 < H < 1$, and for any sample $\omega \in \Omega$, the initial vorticity for (5.9) is not a Radon measure. Consequently, the initial data does not belong to the Delort class and there are no existence results for the corresponding samples. Hence, fractional Brownian motion does not fall within the ambit of any of the available well-posedness theories for two-dimensional Euler equations.
- The Hurst index H in (5.9) controls the regularity (and also roughness) of the initial data (pathwise). Roughly speaking, each sample is Hölder continuous with exponent H. Hence, we can consider a very wide range of scenarios in terms of roughness of the initial data by varying the Hurst-index H, see Fig. 11 for realizations of the horizontal velocity field for three different Hurst indices. In particular, one can observe from this figure that lowering the value of H leads to oscillations of both higher amplitude and frequency in the initial velocity field.

5.4.1. *Structure functions and Compensated energy spectra*

In order to verify convergence of the approximations, generated by Algorithm 4.1, for the fractional Brownian motion initial data (5.9), we will check if the computed structure functions (5.4) decay uniformly with respect to resolution, on decreasing correlation lengths. In Fig. 12 (top row), we plot the structure function at time $T = 1$ for three different Hurst indices of $H = 0.75, 0.5, 0.15$ and observe that the structure functions indeed decay to zero at a certain exponent (independent



(a) $H = 0.15$ (b) $H = 0.5$ (c) $H = 0.75$

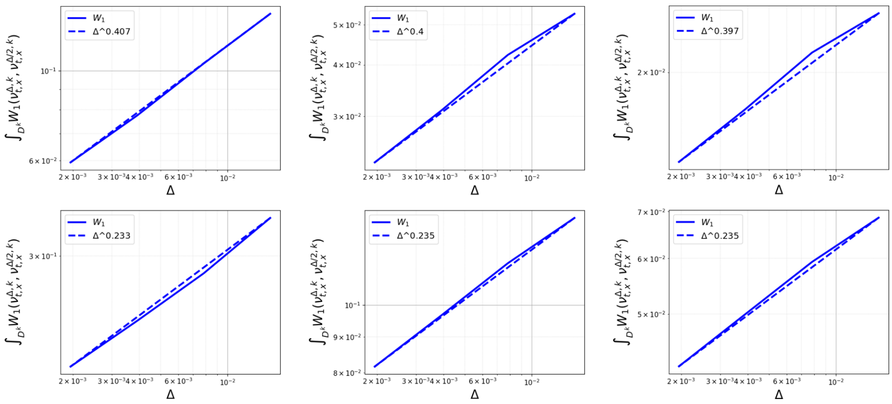
Fig. 12. Instantaneous structure function (5.4) (top row) and compensated energy spectrum (5.6) (bottom row) for fractional Brownian motion initial data with three different Hurst indices at time $T = 1$. The compensated energy spectrum (5.6) is computed with $\gamma = 1.3$ ($H = 0.15$), $\gamma = 2.0$ ($H = 0.5$) and $\gamma = 2.5$ ($H = 0.75$).

of resolution). These exponents are approximately 0.8 for initial $H = 0.75$, 0.6 for the standard Brownian motion initial data ($H = 0.5$) and 0.55 for the initially rough $H = 0.15$. These results indicate that the conditions of the compactness Theorem 2.2 are fulfilled and the approximations converge to a statistical solution of (1.1).

This convergence is further reinforced by the computed compensated energy spectra (5.6), at time $T = 1$, for the three different Hurst indices shown in Fig. 12 (bottom row). Based on the value of the Hurst index, we choose the compensating index $\gamma = 2.5, 2, 1.3$ for the $H = 0.75$, $H = 0.5$, $H = 0.15$, respectively. These values of γ are chosen to provide the correct scaling of the energy spectra at the initial time $t = 0$. As seen from Fig. 12, the compensated energy spectra remain bounded up to the final time $t = T$, independent of the spectral resolution. Hence, the energy spectrum decays at least at the rate of $K^{-\gamma}$ for increasing wave number K, in the inertial range. Consequently, we can readily apply Proposition 4.4 and conclude that the approximations, generated by Algorithm 4.1, converge to a statistical solution, for all three values of the Hurst index H in (5.9).

5.4.2. *Convergence in Wasserstein distance*

Next, we seek to verify convergence of observables (statistical quantities of interest). To this end, we follow the previous section and compute the Wasserstein distances $\int_{D^k} W_1(\nu^{\Delta,k}_{t,x}, \nu^{\Delta/2,k}_{t,x})dx$, corresponding to the k-point correlation marginals for the three different Hurst indices of $H = 0.75, 0.5, 0.15$. In Fig. 13, these metrics are computed at time $T = 1$, for $k = 1, 2$, corresponding to one-point and two-point statistical quantities of interest. As observed from the figure, the approximations



(a) $H = 0.15$ (b) $H = 0.5$ (c) $H = 0.75$

Fig. 13. Wasserstein distances $\int_{D^k} W_1(\nu^{\Delta,k}_{t,x}, \nu^{\Delta/2,k}_{t,x})dx$ for $k = 1$ (top row) and $k = 2$ (bottom row) for fractional Brownian motion initial data with three different Hurst indices at time $t = 1$.

clearly converge in this metric for both one- and two-point statistics, at rates which are independent of the underlying initial Hurst index. The two-point correlation marginals appear to converge at a slower rate than the one-point Young measures. These results validate convergence of all one- and two-point statistical quantities of interest. Taken together with the results for the structure function, compensated energy spectra and Theorem 2.2, they strongly suggest convergence in metric d_T (2.4) on $L^1_T(\mathcal{P})$.

5.4.3. *Comparison between incompressible cases and compressible cases*

Very similar Brownian motion initial data as considered in this section have also been considered to compute statistical solutions for the compressible Euler equations in Ref. 16. While the initial data is indeed very similar, and we observe similar convergence rates in the Wasserstein norm, there appear to be important differences having to do with the different physical processes involved, and in particular shock formation due to compressibility in Ref. 16. Indeed, the time-evolution in the compressible case leads to formation of shock-fronts which sweep through the domain, leaving behind smoother regions in the flow and colliding with each other in the process. In Ref. 16, these shock fronts apparently lead to a dynamic regularization of the flow (as measured by the exponent in the algebraic decay of the structure functions). This regularization is clearly visible in the temporal evolution of the algebraic decay of the structure functions (cf. Fig. 18 in Ref. 16). In contrast, the incompressibility constraint appears to rule out a similar dynamical regularization in the present numerical experiments (at least in the two-dimensional case). Instead, we observe a largely uniform scaling of the energy spectra, without regularization by the fluid dynamics (apart from the expected regularization by numerical diffusion at high wave-numbers).

5.5. *Stability of the computed statistical solution*

The afore-presented numerical experiments clearly validate the convergence theory developed in this paper. Therefore, Algorithm 4.1 provides a practical way to compute statistical solutions. We use this algorithm to study the sensitivity/stability of the computed statistical solution in the context of the discontinuous flat vortex sheet initial data ($\rho = 0$ in (5.2)) to the following perturbations:

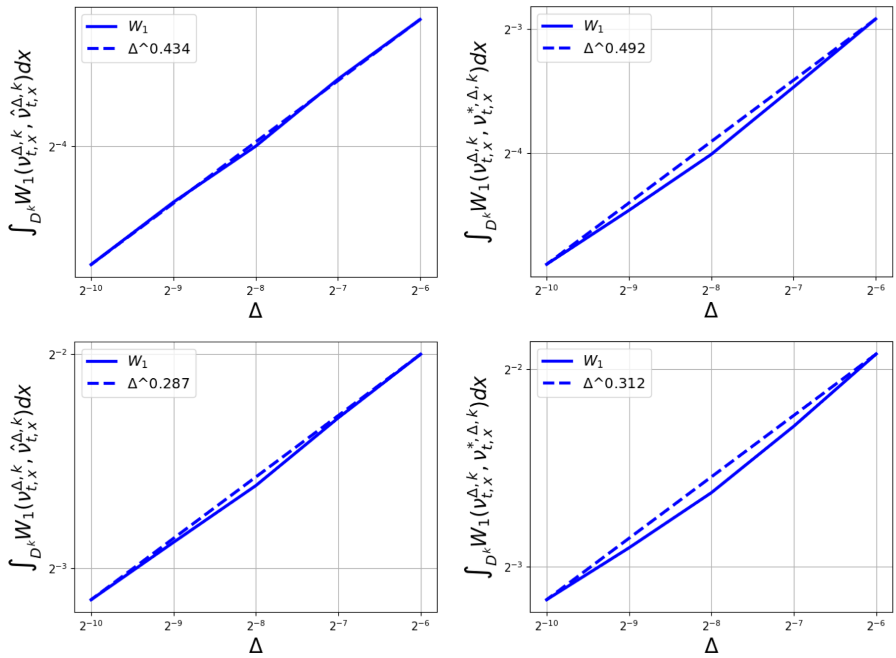
- Type of the initial perturbation, with respect to the underlying probability distribution.
- Type of underlying numerical method.
- Amplitude of the initial perturbation.

We start by changing the type of the underlying initial probability measure in the flat vortex sheet initial data. Instead of choosing the i.i.d. random variables $\alpha_k \sim \mathcal{U}[-1,1]$ according to a uniform distribution as before in (5.2), we choose them from a normal distribution $\alpha_k \sim \mathcal{N}(0, 1/3)$ with mean 0 and variance 1/3. The mean and variance are chosen so the distribution's mean and variance are consistent with those of $\mathcal{U}[-1,1]$.

In order to compare the corresponding approximate statistical solutions, we compute the Wasserstein distance $\int_{D^k} W_1(\nu_{t,x}^{\Delta,k}, \hat{\nu}_{t,x}^{\Delta/2,k})dx$, at different resolutions Δ. Here, $\hat{\mu}^\Delta$ is the statistical solution computed with normally distributed initial data and $\hat{\nu}^\Delta$ is the corresponding correlation measure. We set the initial perturbation amplitude $\delta = 0.05$ in (5.2), $t = 0.4$ and $k = 1, 2$ and plot the computed Wasserstein distances in Fig. 14 (left column). As seen from this figure, the two approximate solutions clearly converge to each other in this metric as the resolution is increased. This indicates that the computed statistical solutions are stable with respect to the variation of underlying initial probability measures.

Next, we consider if the computed statistical solution depends on the underlying numerical method. To this end, we compute the statistical solution for the discontinuous flat vortex sheet initial data (5.2) with Algorithm 4.1 but replace the spectral viscosity method with a finite difference projection method.[3,8,27] The convergence of this method to a statistical solution is considered in a forthcoming thesis.[42] We denote the computed statistical solutions with this method as $\mu_t^{*,\Delta}$ and the corresponding correlation measures as $\nu^{*,\Delta}$ and compute the Wasserstein distance $\int_{D^k} W_1(\nu_{t,x}^{\Delta,k}, \nu_{t,x}^{*,\Delta/2,k})dx$ at time $t = 0.4$ and $k = 1, 2$ and plot the results in Fig. 14 (right column). From this figure, we readily conclude that the statistical solutions computed with the finite difference projection method converge to that computed with the spectral viscosity method on increasing resolution. This strongly suggests the stability of the computed statistical solutions to the underlying (convergent) numerical method.

Finally, we compute approximate statistical solutions for different amplitudes of the initial (random) perturbation in (5.2) by taking different values of the perturbation amplitude ranging from $\delta = 0.05$ to $\delta = 0.05/32$, corresponding to smaller



(a) Type

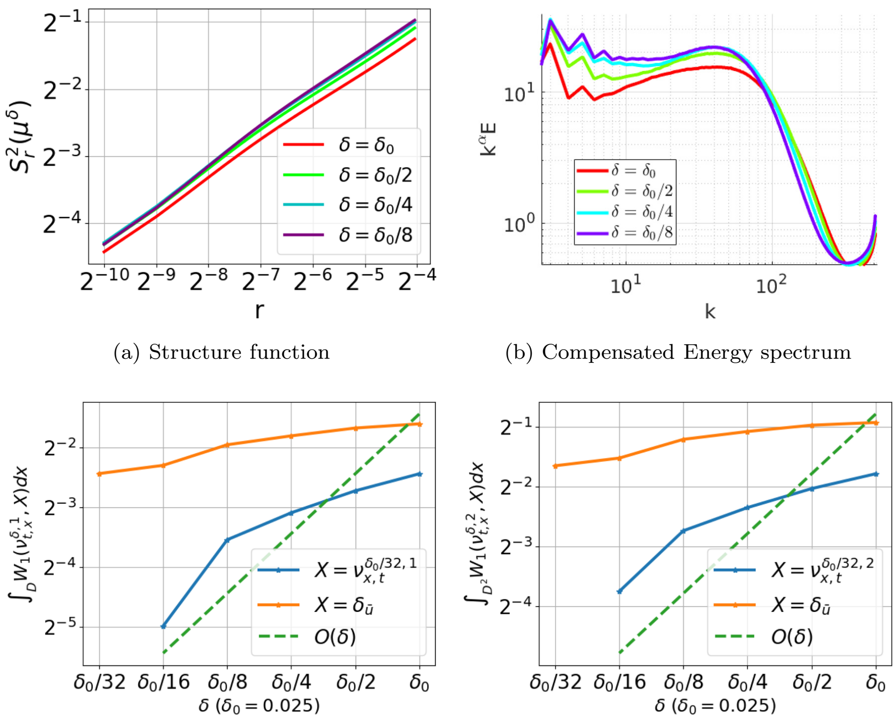
(b) Scheme

Fig. 14. Stability of the flat vortex sheet (5.2) to perturbations. For $k = 1$ (top row) and $k = 2$ (bottom row) and at time $t = 0.4$, we plot left column: Distance $\int_{D^k} W_1(\nu_{t,x}^{\Delta,k}, \hat{\nu}_{t,x}^{\Delta,k})dx$ with $\hat{\nu}$ the correlation marginal corresponding to a normally distributed initial measure in (5.2). Right column: Distance $\int_{D^k} W_1(\nu_{t,x}^{\Delta,k}, \nu_{t,x}^{*,\Delta,k})dx$ with ν^* the correlation marginal computed with the finite difference projection method.

and smaller perturbations of the underlying flat vortex sheet. We computed the statistical solution at the highest resolution of $N = 1024$ with $m = N$ Monte Carlo samples in Algorithm 4.1.

First, we examine if there is convergence of the computed statistical solutions as $\delta \to 0$. To this end, we compute the instantaneous structure function (5.4) at time $T = 0.4$ for different values of δ and plot the results in Fig. 15 (top row, left). Clearly, the computed structure functions are very close and decay with decreasing correlation length with approximately the same exponent. Thus, we can appeal to Theorem 2.2 and claim that the approximate statistical solutions converge. This is further reinforced when we compute the instantaneous compensated energy spectrum (4.13) at time $T = 0.4$ and with $\gamma = 2$. We observe from Fig. 15(top row, right) that the computed energy spectra with different values of δ are very close and decay at the approximately the same rate. Hence, from Theorem 4.4, the underlying approximate statistical solutions will converge.

After establishing convergence of statistical solutions when the amplitude of perturbations in the initial datum (5.2) is decreased, it is natural to ask what these solutions converge to. One possibility is the initial discontinuous flat vortex sheet



(a) Structure function (b) Compensated Energy spectrum

(c) 1-pt Wasserstein distance (d) 2-pt Wasserstein distance

Fig. 15. (Color online) Stability of the flat vortex sheet (5.2) to amplitude of perturbations δ. Top row: Instantaneous structure function (5.4) (left) and compensated energy spectrum (4.13) with $\gamma = 2$ (right) at time $T = 0.4$ for different values of δ. Bottom row: Wasserstein distances with respect to the stationary solution (orange) and reference statistical solution (blue) for the 1-point (left) and 2-point (right) correlation marginals.

itself. After all, setting $\rho, \delta \to 0$ in (5.2), it is easy to see that the discontinuous flat vortex sheet is a stationary solution of the incompressible Euler equations. To check if this stationary solution is indeed the zero perturbation limit of the computed statistical solution, we compute the Wasserstein distances $\int_D W_1(\nu_{t,x}^{\delta,1}, \delta_{\bar{u}(x)})\,dx$ and $\int_{D^2} W_1(\nu_{t,x,y}^{\delta,2}, \delta_{\bar{u}(x)\otimes\bar{u}(y)})\,dx\,dy$ for different values of the perturbation parameter δ with $t = 0.4$ and $k = 1, 2$. Here, $\bar{u}$ is the stationary solution corresponding to the flat vortex sheet initial data, i.e. $\rho, \delta \to 0$. We display these distances in Fig. 15 (bottom row). We observe from this figure that there does not seem to be any perceptible evidence of convergence to the stationary solution. This is consistent with the findings in Ref. 25, where the authors had observed a non-atomic measure-valued solution as the limit of the perturbations to the flat vortex sheet.

In order to further identify the limit of the computed statistical solutions as $\delta \to 0$, we compute a *reference solution* μ_t^{ref}, by setting $\delta = \frac{0.05}{32}$ in (5.2) and $N = m = 1024$ as the spectral resolution and Monte Carlo samples. We compute the Wasserstein distances $\int_{D^k} W_1(\nu_{t,x}^{\delta,k}, \nu_{t,x}^{\mathrm{ref},k})\,dx$ for different values of the perturbation parameter δ with $t = 0.4$ and $k = 1, 2$. Here, ν^{ref} is the correlation measure, corresponding to the reference statistical solution. The distances, plotted in Fig. 15(bottom row), do decrease as δ is decreased, albeit at a slow rate. However, this decrease is clearly visible when compared to the lack of convergence to the stationary solution in the same figures. Thus, we have established that the computed statistical solutions are stable with respect to the amplitude of perturbations and moreover, converge to a probability measure that is different from the one concentrated on the stationary solution.

6. Discussion

We consider the incompressible Euler equations (1.1) in this paper. The existence of classical (or weak) solutions is an outstanding open question in three space dimensions. Although weak solutions are known to exist in two space dimensions, even for very rough initial data, they may not be unique. Similarly, numerical experiments reveal that standard numerical methods may not converge, or converge very slowly, to weak solutions on increasing resolution.

Given these inadequacies of traditional notions of solutions, it is imperative to find solution concepts for (1.1) that are well-posed and amenable to efficient numerical approximation. In this context, we consider the solution framework of statistical solutions in this paper. Statistical solutions are time-parameterized probability measures on $L^2(D; \mathbb{R}^d)$. Given the characterization of probability measures on L^p spaces in Ref. 14, these measures are equivalent to so-called *correlation measures*, i.e. Young measures on tensor-products of the underlying domain and phase space that represent multi-point spatial correlations. Furthermore, we require statistical solutions to satisfy an infinite number of PDEs (see Definition 3.1) for the moments of the underlying correlation measure. Hence, a statistical solution can be interpreted as a measure-valued solution (in the sense of Ref. 11), augmented with information about the evolution of all possible multi-point spatial correlations.

Our aim in this paper was to study the well-posedness and efficient numerical approximation of statistical solutions. To this end, first, we had to characterize convergence on a weak topology on the space $L^1_t(\mathcal{P}(L^2(D; \mathbb{R}^d)))$, under an assumption of *time-regularity* on the underlying measures. Convergence in this topology amounted to convergence of a very large class of observables (or statistical quantities of interest). We then proposed a notion of dissipative statistical solutions and also proved partial well-posedness results for them in a generic sense, namely when the initial measure is concentrated on functions sufficiently near initial data for which smooth solutions exist. This led to short-time well-posedness if the initial probability measure is concentrated on smooth functions. In two space dimensions, we proved global well-posedness for statistical solutions when the initial data is concentrated on smooth functions. Moreover, we also proved a suitable variant of weak–strong uniqueness.

Our main contribution in this paper is the proposal of an Algorithm 4.1 to approximate statistical solutions of the Euler equations. This Monte Carlo type algorithm is a variant of the algorithms proposed recently[13,16,25] and is based on an underlying spectral hyper-viscosity spatial discretization. Under verifiable hypotheses, we prove that the approximations converge in our proposed topology to a statistical solution. These hypotheses either rely on a suitable scaling (or uniform decay) for the structure function, or equivalently, on finding an inertial range (of wave numbers) on which the energy spectrum decays (uniformly in resolution). These hypotheses are very common in the extensive literature on turbulence (see Ref. 20 and references therein). *A key novelty in this paper is the rigorous proof of the fact that easily verifiable conditions on the structure functions or energy spectrum imply a rather strong form of convergence for (multi-point) statistical quantities of interest.* For instance, we observe a surprising fact that a bound on the compensated energy spectrum (5.6) implies that k-point statistics of interest, even for large k, converge. The convergence results also provide a conditional global existence result for statistical solutions in both two and three space dimensions.

We present results of several numerical experiments for the two-dimensional Euler equations. From the numerical experiments, we observe that

- Our convergence theory is validated by all the numerical experiments. The assumptions on the structure functions and energy spectra appear to be very clearly fulfilled in practice. Moreover, the computed solutions converge to a statistical solution in suitable Wasserstein metrics on multi-point correlation marginals. In particular, all admissible observables of interest such as mean, variance, higher moments, structure functions, spectra, multi-point correlation functions, converge on increasing resolution and sample augmentation.
- In clear contrast to the deterministic case where computed solutions may converge very slowly even if one can prove convergence of the underlying numerical method (see Ref. 26 and Fig. 8), statistical quantities of interest seem to be better behaved and converge faster.
- For our numerical examples, we observe convergence of approximations even when the initial data was quite rough such as when the initial vorticity may not have definite sign (as in the flat vortex sheet) or may not even be a Radon measure (as in the fractional Brownian motion with any Hurst index $H \in (0,1)$). For such initial data, the samples are not in the Delort class and the convergence (and existence) theory for two-dimensional Euler equations is no longer valid. On the other hand, we find neat convergence to a statistical solution.
- The computed statistical solutions where observed to be stable with respect to amplitude and type of perturbations of initial data. Moreover, we observed that two very different numerical methods converge to the same statistical solution for a fixed initial data. This is in contrast to the deterministic case, where the computed solutions can differ significantly.[27]

Based on the above discussion, we conclude that statistical solutions are a promising solution framework for the incompressible Euler equations. In particular, there is some scope for proving well-posedness results within this class, possibly with further admissibility criteria. Moreover, numerical approximation of statistical solutions is feasible with ensemble averaging algorithms. Statistical solutions can be a suitable framework for uncertainty quantification and Bayesian inversion for the Euler equations and to encode and explain numerous computational and experimental results for turbulent fluid flows.

There are several limitations in this paper, which provide directions for future work. At the theoretical level, we seek to either relax the criteria on scaling of structure functions or prove it. This will pave the way for global existence results. Similarly, the weak–strong uniqueness results of this paper could be improved.

In terms of numerical approximation, the main issue with the Monte Carlo type Algorithm 4.1 is the slow convergence (in terms of number of samples). This necessitates a very high computational cost, particularly in three space dimensions. We plan to consider efficient variants such as multi-level Monte Carlo,[15,28,40] Quasi-Monte Carlo and deep learning algorithms,[37] for computing statistical solutions of the incompressible Euler equations in three space dimensions, in forthcoming papers.

Appendix A. Proof of Compactness, Theorem 2.1

We first state the following uniform approximation principle, whose proof is an exercise in general topology.

Lemma A.1. *Let* (X, d) *be a complete metric space with metric* d*. Let* $K \subset X$*. If for any* $\epsilon > 0$, *there exists a mapping* $i_\epsilon : K \to X$, *such that the image* $i_\epsilon(K) \subset X$ *is precompact* (*i.e. has compact closure*), *and* $d(x, i_\epsilon(x)) < \epsilon$ *for all* $x \in K$, *then* K *is precompact.*

Proof of Theorem 2.1. We shall only use the implication (2) $\Rightarrow$ (1) in the present work. We prove this result here. For a proof of the converse, (1) $\Rightarrow$ (2), we refer instead to Ref. 24.

Step 1: Fix a mollifier ρ_ϵ for $\epsilon > 0$. Define $i_\epsilon : \mathcal{P}(L^2_x) \to \mathcal{P}(L^2_x)$ by mollification against ρ_ϵ,

$$u_\epsilon(x) := (i_\epsilon u)(x) := \int_D \rho_\epsilon(y) u(x-y)\, \mathrm{d}y.$$

Note that for any fixed $\epsilon > 0$, the image $i_\epsilon(B_M(0))$ is precompact in L^2_x. Let us denote by $i_\epsilon\# : \mathcal{P}(L^2_x) \to \mathcal{P}(L^2_x)$ the push-forward mapping associated with the mollification map i_ϵ. Note that, since all of the probability measures $\mu \in \mathcal{F}$ are supported on $\overline{B_M(0)} \subset L^2_x$, the corresponding push-forward measures $\mu^\epsilon := i_\epsilon\#\mu$ have support in the (compact) set $\overline{i_\epsilon(B_M(0))} \subset L^2_x$. In particular, this implies that, for fixed $\epsilon > 0$, the family $\mathcal{F}_\epsilon := \{\mu^\epsilon; \mu \in \mathcal{F}\}$ is tight. By Prokhorov's theorem, this implies that $\mathcal{F}_\epsilon$ is precompact in the weak topology.

Step 2: We claim that there exists a constant $C > 0$ (independent of $\mathcal{F}$ and $\epsilon > 0$), such that

$$W_1(\mu^\epsilon, \mu) \le C\omega(\epsilon), \tag{A.1}$$

for all $\mu \in \mathcal{F}$. Here, we recall that $\omega(r)$ is the uniform modulus of continuity which exists by assumption on $\mathcal{F}$.

To see (A.1), we note that given any Lipschitz continuous function $F : L^2_x \to \mathbb{R}$ with $\|F\|_{\mathrm{Lip}} \le 1$, and $\mu \in \mathcal{F}$, we have

$$\begin{aligned}
\int_{L^2_x} F(u)\,\mathrm{d}\mu^\epsilon - \int_{L^2_x} F(u)\,\mathrm{d}\mu &= \int_{L^2_x} [F(u_\epsilon) - F(u)]\,\mathrm{d}\mu \\
&\overset{(\|F\|_{\mathrm{Lip}}\le 1)}{\le} \int_{L^2_x} \|u_\epsilon - u\|_{L^2_x}\,\mathrm{d}\mu \\
&\overset{(\text{Jensen})}{\le} \left(\int_{L^2_x} \|u_\epsilon - u\|^2_{L^2_x}\,\mathrm{d}\mu\right)^{1/2} \\
&= \left(\int_{L^2_x}\int_D |u_\epsilon(x) - u(x)|^2\,\mathrm{d}x\,\mathrm{d}\mu\right)^{1/2}.
\end{aligned}$$

The last integral can be bounded using

$$\begin{aligned}
|u_\epsilon(x) - u(x)|^2 &\le \int_{B_\epsilon(0)} \rho_\epsilon(h)|u(x+h) - u(x)|^2\,\mathrm{d}h \\
&\le C^2 \fint_{B_\epsilon(0)} |u(x+h) - u(x)|^2\,\mathrm{d}h,
\end{aligned}$$

where $C^2 = \|\rho\|_{L^\infty}$ depends only on the (fixed) choice of mollifier ρ. Thus,

$$\int_D |u_\epsilon - u|^2\,dx \le C^2 \int_D \fint_{B_\epsilon(0)} |u(x+h) - u(x)|^2\,dh\,dx. \tag{A.2}$$

It follows from the uniform modulus of continuity assumption on $\mathcal{F}$ that

$$\int_{L^p} F(u)\mathrm{d}\mu^\epsilon - \int_{L^p} F(u)\mathrm{d}\mu \le C\omega(\epsilon),$$

for all $F : L^2_x \to \mathbb{R}$, with $\|F\|_{\mathrm{Lip}} \le 1$. Taking the supremum over all such F on the left and recalling the duality formula for the 1-Wasserstein distance W_1, Eq. (2.3), we recover the claimed estimate (A.1).

Step 3: To conclude the proof of this theorem, we note that by Steps 1 and 2, $\mathcal{F}$ is uniformly approximated by precompact sets in the sense of Lemma A.1 (with $X = \mathcal{P}(L^2_x)$ under the 1-Wasserstein distance, $K = \mathcal{F}$ and $i_\epsilon\# : K \to X$ the pushforward induced by mollification), and hence is itself precompact. □

Appendix B. Proof of Completeness of $L_t^1(\mathcal{P})$, Proposition 2.1

Proof of Proposition 2.1. We wish to prove completeness of the metric space $(L_t^1(\mathcal{P}), d_T)$. Let μ_t^n be a Cauchy sequence, i.e.

$$d_T(\mu_t^n, \mu_t^m) = \int_0^T W_1(\mu_t^n, \mu_t^m)\, dt \to 0, \quad \text{as } m, n \to \infty.$$

It will suffice to prove that there exists a subsequence $\mu_t^{n_j}$ with a limit μ_t, since then

$$\limsup_{n\to\infty} d_T(\mu_t^n, \mu_t) \le \limsup_{n,j\to\infty} d_T(\mu_t^n, \mu_t^{n_j}) = 0,$$

by the Cauchy property.

We now choose a suitable subsequence (which we shall immediately reindex), by requiring that $d_T(\mu_t^{n+1}, \mu_t^n) \le 2^{-n}$, for all $n \in \mathbb{N}$. It follows that $\sum_{n\in\mathbb{N}} d_T(\mu_t^{n+1}, \mu_t^n) \le 1$, and thus

$$\sum_{n=1}^{\infty} W_1(\mu_t^{n+1}, \mu_t^n)$$

is a convergent series in $L^1([0,T))$. In particular, this implies that

$$\sum_{n=1}^{\infty} W_1(\mu_t^{n+1}, \mu_t^n) < \infty, \quad \text{for a.e. } t \in [0,T).$$

For a.e. $t \in [0,T)$, we thus have that μ_t^n is Cauchy in $\mathcal{P}(L_x^p)$. Let μ_t denote the limit, which is defined a.e. on $[0,T)$. To see that $t \mapsto \mu_t$ is weak-$*$ measurable, we note that for any $F \in C_b(L_x^p)$, we have that $t \mapsto \langle \mu_t^n, F\rangle$ is measurable, and converges almost everywhere to $\langle \mu_t, F\rangle$. It follows that also $t \mapsto \langle \mu_t, F\rangle$ is measurable. By definition, this means that $t \mapsto \mu_t$ is weak-$*$ measurable.

To show that $\mu_t^n \to \mu_t$ in $L_t^1(\mathcal{P})$, we note that for almost every $t \in [0,T)$, we have

$$\begin{aligned} W_1(\mu_t^n, \mu_t) &= \lim_{m\to\infty} W_1(\mu_t^n, \mu_t^m) \\ &\le \lim_{m\to\infty} \sum_{k=n}^{m} W_1(\mu_t^{k+1}, \mu_t^k) \\ &= \sum_{k=n}^{\infty} W_1(\mu_t^{k+1}, \mu_t^k). \end{aligned}$$

Integrating over $[0,T)$, it follows that

$$d_T(\mu_t^n, \mu_t) \le \sum_{k=n}^{\infty} d_T(\mu_t^{k+1}, \mu_t^k) \le 2^{-n} \to 0, \quad \text{as } n \to \infty.$$

□

Appendix C. Proof of Time-Compactness, Theorem 2.2

Lemma C.1. *If $\mu_t^\Delta \in L_t^1(\mathcal{P})$ is a uniformly time-regular family for which there exists a modulus of continuity $\omega(r)$, such that*

$$S_r^2(\mu_t^\Delta, T) \le \omega(r), \quad \forall\,\Delta > 0$$

and if there exists $M > 0$, such that

$$\int_{L_x^2} \|u\|_{L_x^2}\, d\mu_t^\Delta(u) \le M, \quad \text{for a.e. } t \in [0,T), \quad \forall\,\Delta > 0,$$

then $t \mapsto \mu_t^\Delta$ is L^1-equicontinuous, in the sense that for any $\epsilon > 0$, there exists $\delta > 0$, such that if $h < \delta$, then

$$\int_0^{T-h} W_1(\mu_{t+h}^\Delta, \mu_t^\Delta)\, dt < \epsilon, \quad \forall\ \Delta > 0,$$

Proof. Since μ_t^Δ is uniformly time-regular, by definition, there exists $L > 0$ and $C > 0$, and transfer plans $\pi_{s,t}^\Delta$ between μ_s^Δ and μ_t^Δ, such that

$$\int_{L_x^2 \times L_x^2} \|u - v\|_{H_x^{-L}}\, d\pi_{s,t}^\Delta(u,v) \le C|t-s|.$$

We note that by definition of W_1 as the infimum over all transport plans, we have

$$W_1(\mu_{t+h}^\Delta, \mu_t^\Delta) \le \int_{L_x^2 \times L_x^2} \|u-v\|_{L_x^2}\, d\pi_{t+h,t}^\Delta(u,v).$$

Integrating in time, we thus find

$$\int_0^{T-h} W_1(\mu_{t+h}^\Delta, \mu_t^\Delta)\, dt \le \int_0^{T-h} \int_{L_x^2 \times L_x^2} \|u-v\|_{L_x^2}\, d\pi_{t+h,t}^\Delta(u,v)\, dt.$$

Our goal is to show that the last term converges to 0 as $h \to 0$, uniformly for all $\Delta > 0$. Fix $\epsilon > 0$, arbitrary. We want to find $\delta > 0$, such that $h < \delta$ implies

$$\int_0^{T-h} \int_{L_x^2 \times L_x^2} \|u-v\|_{L_x^2}\, d\pi_{t+h,t}^\Delta(u,v)\, dt < \epsilon, \quad \forall\,\Delta > 0. \tag{C.1}$$

Given $u \in L_x^2$, let $u_\eta := \rho_\eta * u$ denote mollification with a standard mollifier $\rho_\eta(x) = \eta^{-d}\rho(\eta^{-1}x) \ge 0$, supported on a ball B_η of radius η around the origin. Then for all $u, v \in L_x^2$,

$$\|u-v\|_{L_x^2} \le \|u-u_\eta\|_{L_x^2} + \|u_\eta - v_\eta\|_{L_x^2} + \|v_\eta - v\|_{L_x^2},$$

together with the fact that $\mathrm{proj}_1\#\pi_{t+h,t}^\Delta(u,v) = \mu_{t+h}^\Delta(u)$ and $\mathrm{proj}_2\#\pi_{t+h,t}^\Delta(u,v) = \mu_t^\Delta(v)$, implies that

$$\begin{aligned}\int_0^{T-h} &\int_{L_x^2 \times L_x^2} \|u-v\|_{L_x^2}\, d\pi_{t+h,t}^\Delta(u,v)\, dt \\ &\le \int_0^{T-h} \int_{L_x^2} \|u-u_\eta\|_{L_x^2}\, d\mu_{t+h}\, dt\end{aligned}$$

$$
\begin{aligned}
&+\int_0^{T-h}\int_{L_x^2\times L_x^2}\|u_\eta - v_\eta\|_{L_x^2}\,d\pi^\Delta_{t+h,t}\,dt\\
&+\int_0^{T-h}\int_{L_x^2}\|v - v_\eta\|_{L_x^2}\,d\mu_t\,dt\\
\le{}& 2\int_0^T\int_{L_x^2}\|u-u_\eta\|_{L_x^2}\,d\mu_t(u)\,dt\\
&+\int_0^{T-h}\int_{L_x^2\times L_x^2}\|u_\eta - v_\eta\|_{L_x^2}\,d\pi^\Delta_{t+h,t}\,dt.
\end{aligned}
$$

Using the decay of the structure function, we have for all $\Delta > 0$:

$$
\begin{aligned}
\int_0^T\int_{L_x^2}\|u-u_\eta\|_{L_x^2}\,d\mu_t(u)\,dt &\le C'\int_0^T S_\eta^2(\mu_t^\Delta)\,dt\\
&= C'\sqrt{T}S_r^2(\mu_t^\Delta, T) \le C'\sqrt{T}\omega(\eta),
\end{aligned}
$$

where the constant $C' = \|\rho\|_{L^\infty}$. Next, we note that since $H_x^1 \overset{\text{comp}}{\hookrightarrow} L_x^2 \hookrightarrow H_x^{-L}$, it follows from a result originally due to Lions (cf. p. 59 in Ref. 31, Lemma 5.1, p. 58 in Ref. 32, or Lemma 8, p. 84 in Ref. 47) that for any given $\eta > 0$, there exists a constant $Q(\eta) > 0$ such that

$$
\|u\|_{L_x^2} \le \eta^2\|u\|_{H_x^1} + Q(\eta)\|u\|_{H_x^{-L}}, \quad \forall\, u \in H_x^1. \tag{C.2}
$$

Thus, using also that $\|u_\eta - v_\eta\|_{H_x^1} \le C\eta^{-1}\|u-v\|_{L_x^2}$, we can estimate

$$
\begin{aligned}
\|u_\eta - v_\eta\|_{L_x^2} &\le \eta^2\|u_\eta - v_\eta\|_{H_x^1} + Q(\eta)\|u_\eta - v_\eta\|_{H_x^{-L}}\\
&\le C\eta\|u-v\|_{L_x^2} + Q(\eta)\|u-v\|_{H_x^{-L}}\\
&\le C\eta(\|u\|_{L_x^2} + \|v\|_{L_x^2}) + Q(\eta)\|u-v\|_{H_x^{-L}}.
\end{aligned}
$$

Using this estimate, and the uniform time-regularity, we find

$$
\begin{aligned}
&\int_0^{T-h}\int_{L_x^2\times L_x^2}\|u_\eta - v_\eta\|_{L_x^2}\,d\pi^\Delta_{t+h,t}\,dt\\
&\quad\le 2C\eta\int_0^T\left(\int_{L_x^2}\|u\|_{L_x^2}\,d\mu_t^\Delta(u)\right)dt + CQ(\eta)h
\end{aligned}
$$

uniformly in $\Delta > 0$. By assumption, the first term on the right-hand side is bounded by $2C\eta TM$.

Combining these estimates, we find

$$
\int_0^{T-h}\int_{L_x^2\times L_x^2}\|u-v\|_{L_x^2}\,d\pi^\Delta_{t+h,t}\,dt \le C'\omega(\eta) + 2C\eta TM + CQ(\eta)h.
$$

Given $\epsilon > 0$, we can first choose $\eta > 0$ sufficiently small so that $C'\omega(\eta) + 2\eta TM < \epsilon/2$. This fixes a value of $Q = Q(\eta)$ according to Lions' estimate (C.2). Next, we set $\delta := \epsilon/(2CQ)$. It then follows that for $h < \delta$, we have

$$\int_0^{T-h} \int_{L_x^2 \times L_x^2} \|u - v\|_{L_x^2}\, d\pi_{t+h,t}^{\Delta}\, dt < \epsilon, \quad \forall\, \Delta > 0.$$

Thus, given $\epsilon > 0$, we have found $\delta > 0$, such that for all $h < \delta$, we have (C.1). This is what we set out to prove, and concludes our proof of the L^1-equicontinuity of μ_t^{Δ}. □

Next, we note that for any $0 < a < T$, we can associate to any $\mu_t \in L^1([0,T);\mathcal{P})$ a $M_a\mu_t \in L^1([0,T-a];\mathcal{P})$, by defining

$$M_a\mu_t = \frac{1}{a}\int_0^a \mu_{t+h}\, dh. \tag{C.3}$$

The expression on the right-hand side is a straight-forward adaption of the Bochner integral for $L_t^1(\mathcal{P})$.[24] It is then easy to show that

$$W_1(M_a\mu_t, \mu_t) \le \frac{1}{a}\int_0^a W_1(\mu_{t+h}, \mu_t)\, dh.$$

Fix $T_1 < T$. Integrating in time over $[0, T_1]$, it follows that

$$d_{T_1}(M_a\mu_t, \mu_t) \le \frac{1}{a}\int_0^a d_{T_1}(\mu_{t+h}, \mu_t)\, dh, \tag{C.4}$$

for all sufficiently small $a > 0$. In particular, the following lemma is now immediate from this estimate and Lemma C.1.

Lemma C.2. *If μ_t^{Δ} is a equicontinuous family, as in the conclusion of Lemma* C.1, *then for any $\epsilon > 0$, we can find $\delta > 0$, such that for any $a < \delta$, we have*

$$d_{T_1}(M_a\mu_t^{\Delta}, \mu_t^{\Delta}) < \epsilon, \quad \forall\, \Delta > 0.$$

Thus, the family $\{M_a\mu_t^{\Delta}\}_{\Delta>0}$ is a uniform approximation of $\{\mu_t^{\Delta}\}_{\Delta>0}$, in this case.

To show that $\{\mu_t^{\Delta}\}_{\Delta>0}$ is relatively compact in $L^1([0,T_1];\mathcal{P})$, it will thus suffice to prove that $\{M_a\mu_t^{\Delta}\}_{\Delta>0}$ is relatively compact, for any given (fixed) $a > 0$. This is a consequence of the following lemmas and the Arzelà–Ascoli theorem.

Lemma C.3. *Let $T_1 < T$, and fix $0 < a < T - T_1$. Under the assumptions of Theorem* 2.2, *the family $M_a\mu_t^{\Delta}$ is equicontinuous with respect to the W_1-norm, i.e. for any $\epsilon > 0$, there exists $\delta > 0$, such that*

$$W_1(M_a\mu_t^{\Delta}, M_a\mu_s^{\Delta}) < \epsilon, \quad \text{if } |s-t| < \delta.$$

Proof. Fix $a > 0$. Let $\epsilon > 0$ be given. By assumption on the uniform decay of the structure functions $S_r^2(\mu_t^{\Delta}, T)$, it follows from Lemma C.1 that the μ_t^{Δ} are L^1-equicontinuous. Thus, we can find $\delta > 0$, such that

$$\int_0^{T-h} W_1(\mu_{t+h}^{\Delta}, \mu_t^{\Delta})\, dt < a\,\epsilon, \quad \forall\, \Delta > 0,$$

whenever $h < \delta$. Given $h < \delta$, we note that for any 1-Lipschitz continuous $\Phi : L^2_x \to \mathbb{R}$, we have

$$\begin{aligned}\int_{L^2_x} \Phi(u) d(M_a\mu^\Delta_{t+h} - M_a\mu^\Delta_t) &= \frac{1}{a}\int_0^a \int_{L^2_x} \Phi(u)\, d(\mu_{t+h+s} - \mu_{t+s})\, dt \\ &\le \frac{1}{a}\int_0^a W_1(\mu_{t+h+s}, \mu_{t+s})\, ds \\ &\le \frac{1}{a}\int_0^{T-h} W_1(\mu_{h+s}, \mu_s)\, ds \\ &< \epsilon.\end{aligned}$$

Taking the supremum over all 1-Lipschitz continuous Φ, we conclude that

$$W_1(M_a\mu^\Delta_{t+h}, M_a\mu^\Delta_t) \le \epsilon, \quad \forall\, \Delta > 0,$$

whenever $h < \delta$. Thus, $t \mapsto M_a\mu^\Delta_t$ is equicontinuous as an element of $C([0, T - h]; \mathcal{P}(L^2_x))$. □

Lemma C.4. *Let $T_1 < T$, and fix $0 < a < T - T_1$. Under the assumptions of Theorem 2.2, and for any $t \in [0, T_1]$, we have that*

$$\{M_a\mu^\Delta_t \,|\, \Delta > 0\} \subset P(L^2_x)$$

is relatively compact.

Proof. We note that for any $\tau \in [0, T_1]$, we have

$$\begin{aligned} S^2_r(M_a\mu^\Delta_\tau)^2 &= \int_{L^2_x}\int_D \fint_{B_r(0)} |u(x+h) - u(x)|^2\, dh\, dx\, d(M_a\mu^\Delta_\tau) \\ &= \frac{1}{a}\int_0^a \int_{L^2_x}\int_D \fint_{B_r(0)} |u(x+h) - u(x)|^2\, dh\, dx\, d\mu^\Delta_{\tau+s}\, ds \\ &\le \frac{1}{a}\int_0^T \int_{L^2_x}\int_D \fint_{B_r(0)} |u(x+h) - u(x)|^2\, dh\, dx\, d\mu^\Delta_t\, dt \\ &= \frac{1}{a} S^2_r(\mu^\Delta_t, T)^2. \end{aligned}$$

By assumption, there exists a modulus of continuity $\omega(r)$, such that we have $S^2_r(\mu^\Delta_t, T) \le \omega(r)$, uniformly for all Δ. It follows that

$$S^2_r(M_a\mu^\Delta_\tau) \le \frac{1}{\sqrt{a}}\omega(r).$$

Furthermore, it is clear from the definition of $M_a\mu^\Delta_\tau$ (cf. Eq. (C.3)), that if $\mu^\Delta_t(B_M) = 1$ for almost all $t \in [0, T)$, then also $M_a\mu^\Delta_\tau(B_M) = 1$. By Theorem 2.1, it now follows that the family $\{M_a\mu^\Delta_\tau \,|\, \Delta > 0\}$ is pre-compact in $\mathcal{P}(L^2_x)$. □

We are now in a position to prove Theorem 2.2.

Proof of Theorem 2.2. By Lemma C.1, the family μ_t^Δ is equicontinuous in $L^1_x(\mathcal{P})$ (in the sense made precise in Lemma C.1). Then by Lemma C.2, the mollifications $\{M_a\mu_t^\Delta \,|\, \Delta > 0\}_{a>0}$ form a uniform approximation of $\{\mu_t^\Delta \,|\, \Delta > 0\}_{a>0}$ in the sense of Lemma A.1. By the uniform approximation Lemma A.1, to show that $\{\mu_t^\Delta \,|\, \Delta > 0\}$ is precompact it therefore suffices to show that for $a > 0$ fixed, the set $\{M_a\mu_t^\Delta \,|\, \Delta > 0\}$ is precompact. As was shown in Lemmas C.3 and C.4, the mollifications $t \mapsto M_a\mu_t^\Delta$ satisfy:

- The mappings $t \mapsto M_a\mu_t^\Delta$ are (pointwise) equicontinuous with respect to $\Delta > 0$, with values in the metric space $\mathcal{P}(L^2_x)$,
- For each $t \in [0, T_1]$, the set
$$\{M_a\mu_t^\Delta \,|\, \Delta > 0\} \subset \mathcal{P}(L^2_x)$$
is precompact.

Since $\mathcal{P}(L^2_x)$ is a complete metric space under the W_1-metric, it follows from the Arzelà–Ascoli characterization of compact subsets of $C([0, T_1]; \mathcal{P}(L^2_x))$ that $\{M_a\mu_t^\Delta \,|\, \Delta > 0\}$ is precompact in $C([0, T_1]; \mathcal{P}(L^2_x))$, and hence in particular that $\{M_a\mu_t^\Delta \,|\, \Delta > 0\} \subset L^1([0, T_1]; \mathcal{P})$ is precompact. By the uniform approximation lemma, we conclude that we also have that

$$\{\mu_t^\Delta \,|\, \Delta > 0\} \subset L^1([0, T_1]; \mathcal{P})$$

is precompact for any $T_1 < T$.

To conclude the proof, we note that the same argument also applies when time is reversed (or defining the regularization $M_a\mu_t^\Delta$ by averaging over the interval $[t - a, t]$, rather than over $[t, t + a]$). This implies that also

$$\{\mu_t^\Delta \,|\, \Delta > 0\} \subset L^1([T - T_1, T]; \mathcal{P})$$

is precompact for any $T_1 < T$. Combining these results (e.g. setting $T_1 = T/2$), we finally conclude that

$$\{\mu_t^\Delta \,|\, \Delta > 0\} \subset L^1([0, T); \mathcal{P})$$

is precompact. □

Appendix D. Proof of Proposition 2.2

The following lemma provides an estimate for the expected error that is introduced when mollifying random samples from a time-dependent probability measure:

Lemma D.1. *There exists $C > 0$, such that if $\mu_t \in L^1_t(\mathcal{P})$ has structure function $S^2_r(\mu_t, T)$, bounded by a modulus of continuity $\omega(r)$, i.e.*

$$S^2_r(\mu_t, T)^2 \le \omega(r),$$

then for any $\epsilon > 0$, *we have*

$$\int_0^T \int_{L_x^2} \int_D |u_\epsilon - u|^2 \, dx \, d\mu_t(u) \, dt \leq C\omega(\epsilon).$$

Proof. Follows directly from estimate (A.2). □

Proof of Proposition 2.2. To prove Proposition 2.2 and according to the definition of time-regularity Definition 2.2, we need to show that there exists a map $(s,t) \mapsto \pi_{s,t} \in \mathcal{P}(L_x^2 \times L_x^2)$, which defines a transport plan from μ_s to μ_t, and satisfies the H^{-L}-estimate of Definition 2.2. A proof of existence of such $\pi_{s,t}$ will be achieved by viewing each $\pi_{s,t}^\Delta$ as an element of $\mathcal{P}(H_x^{-L} \times H_x^{-L})$, and observing that the $\pi_{s,t}^\Delta$ are concentrated on the compact subset $B_M(0) \times B_M(0) \subset H_x^{-L} \times H_x^{-L}$ for almost all s,t. Let us now choose a sequence $\Delta_k \to 0$. By assumption, for each $k \in \mathbb{N}$, there exists a subset $I_k \subset [0,T)$ of full Lebesgue measure, such that $\pi_{s,t}^{\Delta_k}$ is concentrated on $B_M(0) \times B_M(0) \subset H_x^{-L} \times H_x^{-L}$ for all $s,t \in I_k$, and satisfies

$$\int_{L_x^2 \times L_x^2} \|u - v\|_{H_x^{-L}} \, d\pi_{s,t}^{\Delta_k}(u,v) \leq C|t-s|. \tag{D.1}$$

Let $I = \bigcap_{k=1}^\infty I_k$. Then $I \subset [0,T)$ is of full Lebesgue measure, and $\pi_{s,t}^{\Delta_k}$ is concentrated on $B_M(0) \times B_M(0) \subset H_x^{-L} \times H_x^{-L}$ for all $s,t \in I$ and (D.1) is satisfied for all $s,t \in I$, *uniformly* for all $k \in \mathbb{N}$.

Fix now $s,t \in I$. Since the $\pi_{s,t}^{\Delta_k}$ are concentrated on $B_M(0) \times B_M(0) \subset H_x^{-L} \times H_x^{-L}$, the sequence $\pi_{s,t}^{\Delta_k}$ is tight in $\mathcal{P}(H_x^{-L} \times H_x^{-L})$ (and hence relatively compact under weak convergence by Prokhorov's theorem). We can pass to a weakly convergent subsequence $\pi_{s,t}^{\Delta_{k_j}}$ in $\mathcal{P}(H_x^{-L} \times H_x^{-L})$ such that $\pi_{s,t}^{\Delta_{k_j}} \rightharpoonup \pi_{s,t}$ for some $\pi_{s,t} \in \mathcal{P}(H_x^{-L} \times H_x^{-L})$. Then, by definition of weak convergence, we have for any $F \in C_b(H_x^{-L})$ that

$$\begin{aligned}
\int_{H_x^{-L} \times H_x^{-L}} F(u) \, d\pi_{s,t}(u,v) &= \lim_{j\to\infty} \int_{H_x^{-L} \times H_x^{-L}} F(u) \, d\pi_{s,t}^{\Delta_{k_j}}(u,v) \\
&= \lim_{j\to\infty} \int_{L_x^2 \times L_x^2} F(u) \, d\pi_{s,t}^{\Delta_{k_j}}(u,v) \\
&= \lim_{j\to\infty} \int_{L_x^2} F(u) \, d\mu_s^{\Delta_{k_j}}(u,v) \\
&= \int_{L_x^2} F(u) \, d\mu_s(u).
\end{aligned}$$

When passing to the limit in the last step, we have used the convergence $\mu_s^{\Delta_j} \rightharpoonup \mu_s$ as elements of $\mathcal{P}(L_x^2)$, as well as the fact that if $F \in C_b(H_x^{-L})$, then the restriction $F|_{L_x^2}$ to L_x^2 is in $C_b(L_x^2)$. It follows from $\int F(u) \, d\pi_{s,t}(u,v) = \int F(u) \, d\mu_s(u)$ for all $F \in C_b(H_x^{-L})$ that $\mathrm{proj}_1 \# \pi_{s,t} = \mu_s$. A similar argument shows that also $\mathrm{proj}_2 \# \pi_{s,t} = \mu_t$. In particular, this also implies that $\pi_{s,t}$ is in fact concentrated on $L_x^2 \times L_x^2$.

Indeed, denote $X = L^2_x \subset H^{-L}_x$. Then the complement of the Cartesian product $X \times X$ in $H^{-L}_x \times H^{-L}_x$ satisfies $(X \times X)^c = X^c \times H^{-L}_x \cup H^{-L}_x \times X^c$. This implies that

$$\begin{aligned}\pi_{s,t}((X \times X)^c) &\le \pi_{s,t}(X^c \times H^{-L}_x) + \pi_{s,t}(H^{-L}_x \times X^c)\\ &= \mu_s(X^c) + \mu_t(X^c)\\ &= 0.\end{aligned}$$

In the above estimate, the equality on the second line follows from the fact that $\mathrm{proj}_1\#\pi_{s,t} = \mu_s$ and $\mathrm{proj}_2\#\pi_{s,t} = \mu_t$, and the last equality follows from the fact that μ_s, μ_t are supported on $X = L^2_x \subset H^{-L}_x$.

Furthermore, for any fixed $\gamma > 0$, $(u, v) \mapsto F_\gamma(u, v) = \min(\gamma, \|u - v\|_{H^{-L}_x})$ is a continuous bounded function on $H^{-L}_x \times H^{-L}_x$. We can choose $\gamma > 0$ sufficiently large so that $F_\gamma(u, v) = \|u - v\|_{H^{-L}_x}$ for $(u, v) \in \overline{B_M(0)} \times \overline{B_M(0)}$. It follows that

$$\begin{aligned}\int_{L^2_x \times L^2_x} \|u - v\|_{H^{-L}_x}\, d\pi_{s,t}(u, v) &= \int_{H^{-L}_x \times H^{-L}_x} F_\gamma(u, v)\, d\pi_{s,t}(u, v)\\ &= \lim_{j\to\infty} \int_{H^{-L}_x \times H^{-L}_x} F_\gamma(u, v)\, d\pi^{\Delta_{k_j}}_{s,t}(u, v)\\ &\le \sup_k \int_{L^2_x \times L^2_x} \|u - v\|_{H^{-L}_x}\, d\pi^{\Delta_k}_{s,t}(u, v)\\ &\le C|s - t|.\end{aligned}$$

Since $s, t \in I$ were arbitrary and as $I \subset [0, T)$ is of full Lebesgue measure, we have thus shown that for almost all $s, t \in [0, T)$, there exists a transfer plan $\pi_{s,t}$ from μ_s to μ_t satisfying the assumptions of Definition 2.2. This shows that the limit μ_t is itself time-regular. □

Appendix E. Proof of Theorem 2.4

We now come to the proof of compactness in $L^1_t(\mathcal{P})$, for approximate statistical solutions μ^Δ_t.

Proof of Theorem 2.4. The compactness property follows easily from Theorem 2.2, above. We only provide a proof of strong convergence of the observables here. We drop the subscript j in the following, and assume that $\Delta \to 0$ is a sequence such that $\mu^\Delta_t \to \mu_t$ in $L^1_t(\mathcal{P})$. We recall that by the assumption of this theorem, there exists $M > 0$, such that $\mu^\Delta_t(B_M(0)) = 1$, for all Δ. Fix now $g \in C([0, T) \times D^k \times U^k)$ as in the statement of the theorem. Denote

$$U^\Delta(t, x) := \langle \nu^{\Delta,k}_{t,x}, g(t, x, \xi)\rangle, \quad U(t, x) := \langle \nu^k_{t,x}, g(t, x, \xi)\rangle.$$

By the assumed bound on $g(t, x, \xi)$, we have $U^\Delta, U \in L^1([0, T) \times D^k)$. Given $u \in L^2_x$ and $\epsilon > 0$, let u_ϵ denote the mollification of u. For fixed $\epsilon > 0$, and

$x \in D^k$, the mapping $L^2_x \mapsto U^k$, $u \mapsto u_\epsilon(x)$ is continuous and bounded on $B_M(0) \subset L^2_x$, and so is $u \mapsto g(t,x,u_\epsilon(x))$. Here we recall that in our notation, $u_\epsilon(x) = (u_\epsilon(x_1),\dots,u_\epsilon(x_k)) \in U^k$ for $x = (x_1,\dots,x_k) \in D^k$. We claim that for fixed $\epsilon > 0$, $x \in D^k$, the function

$$F_{x,\epsilon}(u) := g(t,x,u_\epsilon(x)), \quad \text{satisfies } F_{x,\epsilon} \in C_b(\overline{B_M(0)}). \tag{E.1}$$

Indeed, using the assumed Lipschitz bound on $\xi \mapsto g(t,x,\xi)$, we find for $u,v \in \overline{B_M(0)}$ from (2.12), (2.13):

$$\begin{aligned}
&|F_{x,\epsilon}(u) - F_{x,\epsilon}(v)| \\
&\quad = |g(t,x,u_\epsilon(x)) - g(t,x,v_\epsilon(x))| \\
&\quad \le C\sum_{i=1}^k \Pi_i(u_\epsilon(x),v_\epsilon(x))\sqrt{1+|u_\epsilon(x_i)|^2+|v_\epsilon(x_i)|^2}|u_\epsilon(x_i)-v_\epsilon(x_i)|.
\end{aligned}$$

Using the Hölder estimate $|u_\epsilon(x)| \le \|\rho_\epsilon\|_{L^2_x}\|u\|_{L^2_x}$ for the point-values of the mollification, it follows that

$$|F_{x,\epsilon}(u) - F_{x,\epsilon}(v)| \le Ck(|D| + 2M^2\|\rho_\epsilon\|^2_{L^2})^{k-1/2}\|\rho_\epsilon\|_{L^2_x}\|u-v\|_{L^2_x}, \tag{E.2}$$

for all $u,v \in \overline{B_M(0)}$; this implies that $u \mapsto F_{x,\epsilon}(u)$ is Lipschitz-continuous on $\overline{B_M(0)}$, with

$$\|F_{x,\epsilon}\|_{\mathrm{Lip}} \le Ck(|D| + 2M^2\|\rho_\epsilon\|^2_{L^2_x})^{k-1/2}\|\rho_\epsilon\|_{L^2_x}.$$

The previous observation allows us to define

$$U^\Delta_\epsilon(t,x) := \int_{L^2_x} g(t,x,u_\epsilon(x))\,d\mu^\Delta_t(u)$$

and similarly $U_\epsilon(t,x)$ for μ_t. Then, we have for any $\epsilon > 0$ and $\phi \in L^\infty([0,T) \times D^k)$:

$$\begin{aligned}
&\int_0^T\int_{D^k}[U^\Delta_\epsilon(t,x) - U^\Delta(t,x)]\phi(t,x)\,dx\,dt \\
&\quad = \int_0^T\int_{L^2_x}\int_{D^k}[g(t,x,u_\epsilon(x)) - g(t,x,u(x))]\phi(t,x)\,dx\,d\mu^\Delta_t(u)\,dt \\
&\quad \le \int_0^T\int_{L^2_x}\int_{D^k} C\|\phi\|_{L^\infty_{t,x}}\sum_{i=1}^k \Pi_i(u_\epsilon(x),u(x)) \\
&\qquad \times \sqrt{1+|u(x_i)|^2+|u_\epsilon(x_i)|^2}|u_\epsilon(x_i)-u(x_i)|\,dx\,d\mu^\Delta_t(u)\,dt.
\end{aligned}$$

According to Definition (2.13), $\Pi_i(u_\epsilon(x),u(x))$ depends only on $x_1,\dots,x_{i-1},x_{i+1},\dots,x_k$. We can thus integrate over these variables to find

$$\int_{D^{k-1}}\Pi_i(u_\epsilon(x),u(x))\,dx_1\cdots dx_{i-1}\,dx_{i+1}\cdots dx_k \le (|D| + 2M^2)^{k-1},$$

for all u in the support of μ_t^Δ. Where we recall that μ_t^Δ is supported on $\overline{B_M(0)} = \{\|u\|_{L^2_x} \le M\}$, by assumption. On the other hand, from Hölder's inequality, we have

$$\int_D \sqrt{1+|u(x_i)|^2+|u_\epsilon(x_i)|^2}|u_\epsilon(x_i)-u(x_i)|\,dx_i \le (|D|+2M^2)^{1/2}\|u-u_\epsilon\|_{L^2_x},$$

on the support of μ_t^Δ. Combining these estimates, we conclude that

$$\begin{aligned}
&\int_0^T \int_{D^k} \left[U_\epsilon^\Delta(t,x) - U^\Delta(t,x)\right]\phi(t,x)\,dx\,dt \\
&\quad \le C\|\phi\|_{L^\infty_{t,x}} k(|D|+2M^2)^{k-1/2} \int_0^T \int_{L^2_x} \|u-u_\epsilon\|_{L^2_x}\,d\mu_t^\Delta(u)\,dt \\
&\quad \le C\|\phi\|_{L^\infty_{t,x}} k(|D|+2M^2)^{k-1/2}\sqrt{T}\left(\int_0^T \int_{L^2_x} \|u-u_\epsilon\|^2_{L^2_x}\,d\mu_t^\Delta(u)\,dt\right)^{1/2}.
\end{aligned}$$

In the last estimate, we have used Jensen's inequality. Finally, we recall that by (A.2), there exists a constant $C>0$, such that

$$\begin{aligned}
&\left(\int_0^T \int_{L^2_x} \|u-u_\epsilon\|^2_{L^2_x}\,d\mu_t^\Delta(u)\,dt\right)^{1/2} \\
&\quad \le C\left(\int_0^T \int_{L^2_x}\int_D \fint_{B_\epsilon(0)} |u(x+h)-u(x)|^2\,dh\,dx\,d\mu_t^\Delta(u)\,dt\right)^{1/2} \\
&\quad = CS_\epsilon^2(\mu_t^\Delta;T).
\end{aligned}$$

We have thus shown that

$$\int_0^T \int_{D^k} [U_\epsilon^\Delta(t,x)-U^\Delta(t,x)]\phi(t,x)\,dx\,dt \le C\|\phi\|_{L^\infty_{x,t}} S_\epsilon^2(\mu_t^\Delta,T),$$

where $C = C(g,k,M,T)$ depends on the observable $g(t,x,\xi)$, the number of point-correlations k, the uniform *a priori* L^2_x-bound $M>0$, and the maximal time $T>0$, but is independent of Δ, ϵ and ϕ.

Taking the supremum over all ϕ s.t. $\|\phi\|_{L^\infty_{t,x}} \le 1$ on the left-hand side, we obtain

$$\|U_\epsilon^\Delta(t,x)-U^\Delta(t,x)\|_{L^1_{t,x}} \le CS_\epsilon^2(\mu_t^\Delta,T). \tag{E.3}$$

The same inequality holds also true for $U(x,t)$, following the same argument but dropping the subscript Δ in the above estimates. By Lemma D.1, and the uniform (in Δ) upper bound on the structure function, it now follows that there exists an absolute constant $C>0$, such that

$$\|U_\epsilon^\Delta(t,x)-U^\Delta(t,x)\|_{L^1_{t,x}} \le C\omega(\epsilon), \quad \|U_\epsilon(t,x)-U(t,x)\|_{L^1_{t,x}} \le C\omega(\epsilon). \tag{E.4}$$

From the convergence $\mu_t^\Delta \to \mu_t$, it furthermore follows that for any $\epsilon>0$ fixed:

$$\|U_\epsilon^\Delta - U_\epsilon\|_{L^1_{t,x}} \to 0, \quad \text{as } \Delta \to 0. \tag{E.5}$$

Indeed, this follows from the fact that

$$U^{\Delta}(t,x) = \int_{L^2_x} F_{x,\epsilon}(u)\, d\mu_t^{\Delta}(u),$$

where $F_{x,\epsilon}(u)$ has been defined in (E.1) above, and the following estimate:

$$\begin{aligned}
\|U_\epsilon^{\Delta} - U_\epsilon\|_{L^1_{t,x}} &= \int_{D^k}\int_0^T \left|U_\epsilon^{\Delta}(t,x) - U_\epsilon(t,x)\right|\, dt\, dx \\
&= \int_{D^k}\int_0^T \left|\int_{L^2_x} F_{x,\epsilon}(u)[d\mu_t^{\Delta}(u) - d\mu_t(u)]\right|\, dt\, dx \\
&\le \int_{D^k} \left(\int_0^T \|F_{x,\epsilon}\|_{\mathrm{Lip}} W_1(\mu_t^{\Delta}, \mu_t)\, dt\right) dx \\
&\le |D|^k \|F_{x,\epsilon}\|_{\mathrm{Lip}} \left(\int_0^T W_1(\mu_t^{\Delta}, \mu_t)\, dt\right).
\end{aligned}$$

By (E.2), we have $\|F_{x,\epsilon}\|_{\mathrm{Lip}} < \infty$, so that last term converges to 0 as $\Delta \to 0$, as follows from the fact that $\mu_t^{\Delta} \to \mu_t$ in $L^1(\mathcal{P})$.

Thus, combining (E.3)–(E.5), we find

$$\begin{aligned}
\|U^{\Delta}(t,x) - U(t,x)\|_{L^1_{t,x}} &\le \|U^{\Delta}(t,x) - U_\epsilon^{\Delta}(t,x)\|_{L^1_{t,x}} \\
&\quad + \|U_\epsilon^{\Delta}(t,x) - U_\epsilon(t,x)\|_{L^1_{t,x}} \\
&\quad + \|U_\epsilon(t,x) - U(t,x)\|_{L^1_{t,x}} \\
&\le 2C\omega(\epsilon) + \|U_\epsilon^{\Delta}(t,x) - U_\epsilon(t,x)\|_{L^1_{t,x}},
\end{aligned}$$

which implies that

$$\limsup_{\Delta\to 0} \|U^{\Delta}(t,x) - U(t,x)\|_{L^1([0,T)\times D^k)} \le 2C\omega(\epsilon),$$

for any fixed $\epsilon > 0$. Finally, noting that the left-hand side is independent of ϵ, we may let $\epsilon \to 0$ to show that

$$\limsup_{\Delta\to 0} \|U^{\Delta}(t,x) - U(t,x)\|_{L^1([0,T)\times D^k)} \le 0,$$

or, equivalently, that $U^{\Delta}(t,x) \to U(t,x)$ in $L^1([0,T)\times D^k)$, as $\Delta \to 0$. □

Appendix F. Proof of Existence and Uniqueness, Theorem 3.1

We will follow the notation used in Sec. 3.1. Before coming to the proof of Theorem 3.1, we observe that

$$\mathcal{G}_n = \bigcup_{\overline{v}\in\mathcal{C}} B_{r(\overline{v})/n}(\overline{v}), \quad r(\overline{v}) := e^{-C(\overline{v})T}.$$

In particular, $\mathcal{G}_n$ is covered by the open balls $B_{r(\overline{v})/n}(\overline{v})$. Since L^2_x is a separable metric space, we can choose a countable subcovering, i.e. we can find a dense set of initial data $\{\overline{u}_i\}_{i\in\mathbb{N}} \subset \mathcal{C}$ in such a way that

$$\mathcal{G}_n = \bigcup_{i=1}^{\infty} B_{r(\overline{u}_i)/n}(\overline{u}_i)$$

$$= \left\{ \overline{u} \,|\, \exists\, i \in \mathbb{N} \text{ s.t. } \|\overline{u} - \overline{u}_i\|_{L^2_x} < \frac{1}{n} e^{-C(\overline{u}_i)T} \right\}. \tag{F.1}$$

Furthermore, choosing such a countable subset $\{\overline{u}_i^{(n)}\}_{i\in\mathbb{N}}$ for each $\mathcal{G}_n$, $n \in \mathbb{N}$, and considering finally the union $\bigcup_{n=1}^{\infty}\{\overline{u}_i^{(n)}\}_{i\in\mathbb{N}}$, we can in fact find a single countable subset of $\mathcal{C}$ (not depending on n), such that (F.1) holds for all $\mathcal{G}_n$, $n \in \mathbb{N}$.

After this preliminary observation, we now come to the proof of Theorem 3.1.

Proof of Theorem 3.1. Let $\bar{\mu} \in \mathcal{P}(L^2_x)$ be given, such that $\bar{\mu}(\mathcal{G}) = 1$. The idea of the proof is to first choose a countable set $\{\overline{u}_i\}_{i\in\mathbb{N}}$ satisfying (F.1) for all $n \in \mathbb{N}$, and to observe that $\bar{\mu}$ can be well approximated by atomic measures of the form $\bar{\rho} = \sum_{i=1}^{\infty} \alpha_i \delta_{\overline{u}_i}$. We then use the regularity of the solution $t \mapsto u_i(t)$ of the Euler equations with initial data $\overline{u}_i$ to show that a dissipative solution μ_t exists and that it is unique.

Construction of suitable approximants: For any $n \in \mathbb{N}$, let us first construct a suitable approximant $\bar{\rho}^{(n)} \approx \bar{\mu}$, of the form

$$\bar{\rho}^{(n)} = \sum_{i=1}^{\infty} \alpha_i^{(n)} \delta_{\overline{u}_i}. \tag{F.2}$$

We define the coefficients $\alpha_i^{(n)}$, as well as a decomposition $\mathcal{G} = \bigcup_{i=1}^{n} S_i$, recursively as follows: First, denote

$$r_i^{(n)} := \frac{1}{n} e^{-C(\overline{u}_i)T}, \tag{F.3}$$

such that $\mathcal{G}_n = \bigcup_i B_{r_i^{(n)}}(\overline{u}_i)$. Then we set

$$\begin{cases} S_1^{(n)} := B_{r_1^{(n)}}(\overline{u}_1) \cap \mathcal{G}, \\ \Sigma_1^{(n)} := S_1^{(n)}, \\ \alpha_1^{(n)} := \bar{\mu}(S_1^{(n)}) \end{cases} \tag{F.4}$$

and inductively

$$\begin{cases} S_{i+1}^{(n)} := \left[B_{r_{i+1}^{(n)}}(\overline{u}_{i+1}) \backslash \Sigma_i^{(n)} \right] \cap \mathcal{G}, \\ \Sigma_{i+1}^{(n)} := \Sigma_i^{(n)} \cup S_{i+1}^{(n)}, \\ \alpha_{i+1}^{(n)} := \bar{\mu}(S_{i+1}^{(n)}). \end{cases} \tag{F.5}$$

Note that $S_i^{(n)}$ can be thought of as the support of $\bar{\mu}$ close to $\overline{u}_i$, and $\Sigma_i^{(n)}$ keeps track of the set that has already been assigned to $\overline{u}_j$ for some $j < i$. By this construction, we have

$$\begin{aligned} S_i^{(n)} \cap S_j^{(n)} = \emptyset, \quad \text{for } i \neq j, \\ \bigcup_{i=1}^{\infty} S_i^{(n)} = \mathcal{G} \end{aligned} \tag{F.6}$$

and furthermore

$$\sum_{i=1}^{\infty} \alpha_i^{(n)} = \sum_{i=1}^{\infty} \bar{\mu}(S_i^{(n)}) = \bar{\mu}(\mathcal{G}) = 1,$$

for all $n \in \mathbb{N}$. Thus, setting $\bar{\mu}_i^{(n)} := \bar{\mu}|_{S_i^{(n)}}$, it follows that

$$\bar{\mu} = \sum_{i=1}^{\infty} \alpha_i^{(n)} \bar{\mu}_i^{(n)}.$$

Since $\bar{\mu}_i^{(n)}$ is supported in a small neighborhood of $\overline{u}_i$, we expect $\bar{\mu}_i^{(n)} \approx \delta_{\overline{u}_i}$, and it is now natural to define

$$\bar{\rho}^{(n)} := \sum_{i=1}^{\infty} \alpha_i^{(n)} \delta_{\overline{u}_i}.$$

A dissipative statistical solution with initial data $\bar{\rho}^{(n)}$ is given by

$$\rho_t^{(n)} := \sum_{i=1}^{\infty} \alpha_i^{(n)} \delta_{u_i(t)}.$$

In the following, we will show that $(\rho_t^{(n)})_{n=1,2,\ldots}$ is Cauchy for any $t \in [0, T)$, implying the existence of a limit $\rho_t^{(n)} \rightharpoonup \mu_t$. Since the definition of a dissipative statistical solutions is linear in the probability measure, it is not hard to see that such a limit μ_t must itself be a dissipative statistical solution. Furthermore, we will show that the limit μ_t has $\bar{\mu}$ as initial data. Finally, we will show that this μ_t is in fact unique in the class of dissipative statistical solutions.

The sequence $\rho_t^{(n)}$ is Cauchy: Let $m, n \in \mathbb{N}$ be arbitrary. Note that

$$\mathcal{G} = \bigcup_{i=1}^{\infty} S_i^{(m)} = \bigcup_{i=1}^{\infty} S_i^{(n)},$$

implies that upon defining $S_{i,j}^{(m,n)} := S_i^{(m)} \cap S_j^{(n)}$, we have

$$\mathcal{G} = \bigcup_{i,j=1}^{\infty} S_{i,j}^{(m,n)}, \quad S_i^{(m)} = \bigcup_{j=1}^{\infty} S_{i,j}^{(m,n)}, \quad S_j^{(n)} = \bigcup_{i=1}^{\infty} S_{i,j}^{(m,n)}.$$

Furthermore, the $S_{i,j}^{(m,n)}$ are pairwise disjoint by (F.6). Thus, setting

$$\alpha_{i,j}^{(m,n)} := \bar{\mu}(S_{i,j}^{(m,n)}),$$

we can define a transport plan $\pi_t \in \mathcal{P}(L^2_x \times L^2_x)$ from $\rho_t^{(m)} \in \mathcal{P}(L^2_x)$ to $\rho_t^{(n)} \in \mathcal{P}(L^2_x)$, by

$$\pi_t := \sum_{i,j} \alpha_{i,j}^{(m,n)} \delta_{u_i(t)} \otimes \delta_{u_j(t)}.$$

Note that if $S_{i,j}^{(m,n)} \neq \emptyset$, then since $S_i^{(m)} \subset B_{r_i^{(m)}}(\overline{u}_i)$ and $S_j^{(n)} \subset B_{r_j^{(n)}}(\overline{u}_j)$, we must in particular have

$$B_{r_i^{(m)}}(\overline{u}_i) \cap B_{r_j^{(n)}}(\overline{u}_j) \neq \emptyset.$$

This implies that

$$\|\overline{u}_i - \overline{u}_j\|_{L^2_x} \leq r_i^{(m)} + r_j^{(n)} \leq 2\max(r_i^{(m)}, r_j^{(n)}).$$

Assuming now without loss of generality that the maximum is $\max(r_i^{(m)}, r_j^{(n)}) = r_i^{(m)}$, we find from the stability estimate for Lipschitz continuous solutions of the Euler equations that

$$\|u_i(t) - u_j(t)\|_{L^2_x} \leq 2r_i^{(m)} e^{C(u_i)T} = \frac{2}{m} \leq 2\max\left(\frac{1}{m}, \frac{1}{n}\right).$$

It follows that

$$\begin{aligned} W_2(\rho_t^{(m)}, \rho_t^{(n)})^2 &\leq \int_{L^2_x \times L^2_x} \|u - v\|^2_{L^2_x}\, d\pi_t(u,v) \\ &= \sum_{i,j} \alpha_{i,j}^{(m,n)} \|u_i - u_j\|^2_{L^2_x} \\ &\leq 4\max\left(\frac{1}{m}, \frac{1}{n}\right)^2 \sum_{i,j} \alpha_{i,j}^{(m,n)} \\ &= 4\max\left(\frac{1}{m}, \frac{1}{n}\right)^2. \end{aligned}$$

In particular, this shows that $\rho_t^{(n)}$ is Cauchy $\mathcal{P}(L^2_x)$, and has a limit $\rho_t^{(n)} \to \mu_t$.

μ_t is a dissipative statistical solutions with initial data $\bar{\mu}$: To see this, we first note that the 2-Wasserstein distance between $\bar{\rho}^{(n)} = \rho_0^{(n)}$ and $\bar{\mu}$ can be bounded by

$$\begin{aligned} W_2(\bar{\mu}, \bar{\rho}^{(n)})^2 &\leq \sum_{i=1}^{\infty} \alpha_i^{(n)} \int_{L^2_x} \|u - \overline{u}_i\|^2_{L^2_x}\, d\bar{\mu}_i \\ &\leq \sum_{i=1}^{\infty} \alpha_i^{(n)} \int_{L^2_x} [r_i^{(n)}]^2\, d\bar{\mu}_i \leq \sum_{i=1}^{\infty} \alpha_i^{(n)} \frac{1}{n^2} = \frac{1}{n^2}. \end{aligned}$$

Thus, $\rho_0^{(n)}$ converges to $\bar{\mu}$. Since $\rho_t^{(n)} \to \mu_t$ uniformly for $t \in [0, T)$, and each $\rho_t^{(n)}$ is a dissipative statistical solution, this implies that μ_t is a statistical solution with initial data $\bar{\mu}$.

μ_t **is unique:** Let $\widetilde{\mu}_t$ be any dissipative statistical solution. We have to show that $\widetilde{\mu}_t = \mu_t$. To this end, we fix $\epsilon > 0$ for the moment. Since $\rho_t^{(n)} \rightharpoonup \mu_t$, we have

$$W_2(\mu_t, \widetilde{\mu}_t) \le \epsilon + W_2(\widetilde{\mu}_t, \rho_t^{(n)}),$$

for all sufficiently large n. It will thus suffice to show that $W_2(\widetilde{\mu}_t, \rho_t^{(n)})$ can be made smaller than any ϵ for n large. We now fix n, such that $1/n < \epsilon$. We will from now on denote $\alpha_i = \alpha_i^{(n)}$ for $i = 1, 2, \ldots$. Define a finite convex combination of Dirac measures by

$$\widetilde{\rho}_t = \sum_{i=1}^{N} \alpha_i \delta_{u_i(t)} + \alpha_0 \delta_{u_0(t)},$$

where we set $\alpha_0 = \sum_{i>N} \alpha_i$, and $u_0(t) \equiv 0$. Here, N is chosen sufficiently large to ensure that

$$\alpha_0 = \sum_{i>N} \alpha_i < \epsilon^2/M^2,$$

where $M \ge 1$ is chosen such that $\bar{\mu}(B_M^c) = 0$. Such M exists by assumption on $\bar{\mu}$. Note in particular, that for this choice of $\widetilde{\rho}_t$, we have

$$W_2(\widetilde{\rho}_t, \rho_t^{(n)}) \le \epsilon,$$

for all $t \in [0, T)$.

Define now $\bar{\mu}_i \in \Lambda(\alpha, \bar{\mu})$ by

$$\bar{\mu}_i = \bar{\mu}_i^{(n)}, \quad \text{for } i = 1, \ldots, N, \quad \text{and}$$
$$\bar{\mu}_0 = \begin{cases} \dfrac{1}{\alpha_0} \displaystyle\sum_{i>N} \alpha_i \bar{\mu}_i, & \alpha_0 > 0, \\ 0, & \alpha_0 = 0. \end{cases}$$

Corresponding to this decomposition

$$\overline{\mu} = \sum_{i=0}^{N} \overline{\mu}_i$$

and by assumption on the diffusivity of $\widetilde{\mu}_t$ (cf. Definition 3.2), there exists a decomposition $\hat{\mu}_{i,t}$ of $\widetilde{\mu}_t$, such that for each $i = 0, \ldots, N$:

$$\int_0^T \int_{L_x^2} \int_D [u \cdot \partial_t \phi + (u \otimes u) : \nabla \phi] \, dx \, d\hat{\mu}_{i,t}(u) \, dt$$
$$= - \int_{L_x^2} \int_D u \cdot \phi(x, 0) \, dx \, d\overline{\mu}_i(u)$$

and

$$\int_{L_x^2} \|u\|_{L_x^2}^2 \, d\hat{\mu}_{i,t}(u) \le \int_{L_x^2} \|u\|_{L_x^2}^2 \, d\overline{\mu}_i(u).$$

Repeating the argument used in the proof of weak–strong uniqueness for measure-valued solutions,[6] it follows from the fact that $u_i(t)$ is a strong solution, that

$$\int_{L^2_x} \|u - u_i(t)\|^2_{L^2_x} \, d\hat{\mu}_{i,t}(u) \le e^{C(\overline{u}_i)T} \int_{L^2_x} \|u - \overline{u}_i\|^2_{L^2_x} \, d\bar{\mu}_i(u),$$

for almost all $t \in [0, T)$, and all $i = 0, \dots, N$. By our choice of $\bar{\mu}_i$, which is supported in $B_{r_i^{(n)}}(\overline{u}_i)$ and $r_i^{(n)} = e^{-C(\overline{u}_i)T}/n$, this implies

$$W_2(\hat{\mu}_{i,t}, \delta_{u_i(t)})^2 \le \int_{L^2_x} \|u - u_i(t)\|^2_{L^2_x} \, d\hat{\mu}_{i,t}(u) \le \frac{1}{n^2} \le \epsilon^2,$$

for $i = 1, \dots, N$. For $i = 0$, we have instead

$$W_2(\hat{\mu}_{i,t}, \delta_{u_i(t)})^2 \le \int_{L^2_x} \|u - \underbrace{u_i(t)}_{=0}\|^2_{L^2_x} \, d\hat{\mu}_{i,t}(u) \le M^2,$$

by assumption on the L^2-boundedness of $\overline{\mu}$.

We conclude that for almost all $t \in [0, T)$:

$$\begin{aligned} W_2(\widetilde{\mu}_t, \widetilde{\rho}_t)^2 &\le \alpha_0 W_2(\hat{\mu}_{0,t}, \delta_{u_0(t)})^2 + \sum_{i=1}^N \alpha_i W_2(\hat{\mu}_{i,t}, \delta_{u_i(t)})^2 \\ &\le \frac{\epsilon^2}{M^2} M^2 + \sum_{i=1}^N \alpha_i \epsilon^2 \le 2\epsilon^2 \end{aligned}$$

and hence

$$\begin{aligned} W_2(\widetilde{\mu}_t, \mu_t) &\le W_2(\widetilde{\mu}_t, \widetilde{\rho}_t) + W_2(\widetilde{\rho}_t, \rho_t^{(n)}) + W_2(\rho_t^{(n)}, \mu_t) \\ &\le \sqrt{2}\epsilon + \epsilon + \epsilon \le 4\epsilon. \end{aligned}$$

Since $\epsilon > 0$ was arbitrary, it follows that $W_2(\widetilde{\mu}_t, \mu_t) = 0$ for almost all $t \in [0, T)$, i.e. μ_t is the unique dissipative statistical solution with initial data $\bar{\mu}$. □

Acknowledgment

The research of SL and SM is partially supported by the European Research Council Consolidator grant ERC-CoG 770880 COMANFLO.

References

1. R. Abgrall and S. Mishra, Uncertainty quantification for hyperbolic systems of conservation laws, *Handb. Numer. Anal.* **18** (2018) 507–544.
2. C. Bardos and E. Tadmor, Stability and spectral convergence of Fourier method for nonlinear problems: On the shortcomings of the 2/3 de-aliasing method, *Numer. Math.* **129** (2015) 749–782.
3. J. Bell, P. Collela and H. M. Glaz, A second-order projection method for the incompressible Navier–Stokes equations, *J. Comput. Phys.* **85** (1989) 257–283.
4. V. I. Bogachev, *Measure Theory,* Vols. I, II (Springer-Verlag, 2007).

5. N. Bonneel, M. van de Panne, S. Paris and W. Heidrich, Displacement interpolation using Lagrangian mass transport, *ACM Trans. Graph.* **30** (2011) 158:1–158:12.
6. Y. Brenier, C. De Lellis and L. Székelyhidi, Weak-strong uniqueness for measure-valued solutions, *Commun. Math. Phys.* **305** (2011) 351.
7. C. Canuto, M. Y. Hussaini, A. Quarteroni and T. A. Zang, *Spectral Methods: Fundamentals in Single Domains* (Springer Science & Business Media, 2007).
8. A. Chorin, Numerical solution of the Navier–Stokes equations, *Math. Comput.* **22** (1968) 745–762.
9. C. De Lellis and L. Székelyhidi, Jr., The Euler equations as a differential inclusion, *Ann. of Math.* (2) **170** (2009) 1417–1436.
10. J.-M. Delort, Existence de nappes de toubillon en dimension deux, *J. Am. Math. Soc.* **4** (1991) 553–586.
11. R. J. DiPerna and A. J. Majda, Oscillations and concentrations in weak solutions of the incompressible fluid equations, *Commun. Math. Phys.* **108** (1987) 667–689.
12. E. Feireisl, M. Lukáčová-Medviďová and H. Mizerová, $\mathcal{K}$-convergence as a new tool in numerical analysis, *IMA J. Numer. Anal.* **40**(4) (2020) 2227–2255.
13. U. S. Fjordholm, R. Kppeli, S. Mishra and E. Tadmor, Construction of approximate entropy measure valued solutions for hyperbolic systems of conservation laws, *Found. Comput. Math.* **17** (2017) 763–827.
14. U. S. Fjordholm, S. Lanthaler and S. Mishra, Statistical solutions of hyperbolic conservation laws I: Foundations, *Arch. Ration. Mech. Anal.* **226** (2017) 809–849.
15. U. S. Fjordholm, K. O. Lye and S. Mishra, Numerical approximation of statistical solutions of scalar conservation laws, *SIAM J. Numer. Anal.* **56** (2018) 2989–3009.
16. U. S. Fjordholm, K. O. Lye, S. Mishra and F. Weber, Statistical solutions of hyperbolic systems of conservation law: Numerical approximation, *Math. Models Methods Appl. Sci.* **30** (2020) 539–609.
17. U. S. Fjordholm, S. Mishra and E. Tadmor, On the computation of measure-valued solutions, *Acta Numer.* **25** (2016) 567–679.
18. R. Flamary and N. Courty, POT Python Optimal Transport library (2017).
19. C. Foias, R. M. Rosa and R. Temam, A note on statistical solutions of the three-dimensional Navier–Stokes equations: The time-dependent case, *C. R. Math.* **348** (2010) 235–240.
20. U. Frisch, *Turbulence* (Cambridge Univ. Press, 1995).
21. D. Gottlieb and S. A. Orszag, *Numerical Analysis of Spectral Methods: Theory and Applications*, Vol. 26 (SIAM, 1977).
22. J. Guzman, C.-W. Shu and F. A. Sequira, H-Div conforming and DG methods for the incompressible Euler equations, *IMA J. Num. Anal.* **37** (2017) 1733–1771.
23. R. Krasny, A study of singularity formation in a vortex sheet with a point vortex approximation, *J. Fluid Mech.* **167** (1986) 65–93.
24. S. Lanthaler, PhD thesis (2021), in preparation.
25. S. Lanthaler and S. Mishra, Computation of measure-valued solutions for the incompressible Euler equations, *Math. Models Methods Appl. Sci.* **25** (2015) 2043–2088.
26. S. Lanthaler and S. Mishra, On the convergence of the spectral viscosity method for the two-dimensional incompressible Euler equations with rough initial data, *Found. Comput. Math.* **20** (2020) 1309–1362.
27. F. Leonardi, Numerical methods for ensemble based solutions to incompressible flow equations, Ph.D. thesis, ETH Zürich (2018).
28. F. Leonardi, S. Mishra and C. Schwab, Numerical approximation of statistical solutions of planar, incompressible flows., *Math. Model Methods Appl. Sci.* **26** (2016) 2471–2523.

29. P. Lévy, *Processus Stochastiques et Mouvement Brownien*, Les Grands Classiques Gauthier–Villars [Gauthier–Villars Great Classics] (Éditions Jacques Gabay, 1992), followed by a note by M. Loève, Reprint of the second (1965) edition.
30. D. Levy and E. Tadmor, Non-oscillatory central schemes for the incompressible 2-D Euler equations, *Math. Res. Lett.* **4** (1997) 321–340.
31. J. L. Lions, *Equations Differentielles Operationelles* (Springer, 1961).
32. J. L. Lions, *Quelques Méthodes de Résolutions des Problèmes Aux Limites Non Linéaires* (Dunod, 1969).
33. P.-L. Lions, *Mathematical Topics in Fluid Mechanics: Incompressible Models*, Vol. 1, Oxford Lecture Series in Mathematics and its Applications, Vol. 3 (The Clarendon Press, Oxford Univ. Press, 1996).
34. J.-G. Liu and Z. Xin, Convergence of the point vortex method for 2-D vortex sheet, *Math. Comput.* **70** (2001) 595–606.
35. J.-G. Liu and Z. P. Xin, Convergence of vortex methods for weak solutions to the 2-D Euler equations with vortex sheet data, *Commun. Pure Appl. Math.* **48** (1995) 611–628.
36. M. C. Lopes-Filho, H. J. Nussenzveig-Lopes and E. Tadmor, Approximate solutions of the incompressible Euler equations with no concentrations, *Ann. Inst. Henri Poincare (C) NonLinear Anal.* **17** (2000) 371–412.
37. K. O. Lye, S. Mishra and D. Ray, Deep learning observables in computational fluid dynamics, preprint (2019), arXiv:1903.03040v1.
38. A. J. Majda and A. L. Bertozzi, *Vorticity and Incompressible Flow*, Cambridge Texts in Applied Mathematics (Cambridge Univ. Press, 2001).
39. F. Mengual and L. Székelyyhidi Jr., Dissipative Euler flows for vortex sheet initial data without distinguished sign (2020).
40. S. Mishra, C. Schwab and J. Šukys, Multi-level Monte Carlo finite volume methods for nonlinear systems of conservation laws in multi-dimensions, *J. Comput. Phys.* **231** (2012) 3365–3388.
41. S. A. Orszag, Numerical methods for the simulation of turbulence, *Phys. Fluids* **12** (1969) II–250.
42. C. Pares-Pulido, Numerical approximation of statistical solutions of the incompressible Euler equations with a finite difference-projection method (2021), in preparation.
43. V. Scheffer, An inviscid flow with compact support in space-time, *J. Geom. Anal.* **3** (1993) 343–401.
44. A. Schnirelman, Weak solutions with decreasing energy of the incompressible Euler equations, *Commun. Math. Phys* **210** (2000) 541–603.
45. S. Schochet, The rate of convergence of spectral-viscosity methods for periodic scalar conservation laws, *SIAM J. Numer. Anal.* **27** (1990) 1142–1159.
46. Z.-S. She, E. Aurell and U. Frisch., The inviscid Burgers equation with initial data of Brownian type, *Commun. Math. Phys.* **148** (1992) 623–641.
47. J. Simon, Compact sets in the space $L^p(O,T;B)$, *Ann. Mat. Pura Appl.* **146** (1986) 65–96.
48. Y. G. Sinai, Statistics of shocks in solutions of inviscid Burgers equation, *Commun. Math. Phys.* **148** (1992) 601–621.
49. L. Székelyhidi Jr, Weak solutions to the incompressible euler equations with vortex sheet initial data, *C. R. Math.* **349** (2011) 1063–1066.
50. E. Tadmor, Convergence of spectral methods for nonlinear conservation laws, *SIAM J. Numer. Anal.* **26**(1) (1989) 30–44.
51. E. Tadmor, Burgers' equation with vanishing hyper-viscosity, *Commun. Math. Sci.* **2** (2004) 317–324.

52. I. Vecchi and S. Wu, On L1-vorticity for 2-D incompressible flow, *Manuscripta Math.* **78** (1993) 403–412.
53. C. Villani, *Optimal Transport: Old and New*, Vol. 338 (Springer Science & Business Media, 2008).
54. M. I. Vishik and A. V. Fursikov, Homogeneous statistical solutions of the Navier–Stokes system, *Uspehi Mat. Nauk.* **32** (1977) 179–180.