Journal of Nonlinear Science (2021) 31:24
https://doi.org/10.1007/s00332-021-09681-w

Journal of Nonlinear Science

2D Euler Equations with Stratonovich Transport Noise as a Large-Scale Stochastic Model Reduction

Franco Flandoli[1] · Umberto Pappalettera[1]

Received: 3 August 2020 / Accepted: 23 January 2021 / Published online: 11 February 2021

Abstract

The limit from an Euler-type system to the 2D Euler equations with Stratonovich transport noise is investigated. A weak convergence result for the vorticity field and a strong convergence result for the velocity field are proved. Our results aim to provide a stochastic reduction of fluid-dynamics models with three different time scales.

1 Introduction

This work deals with the 2D Euler equations in vorticity form on the two-dimensional torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$:

$$\begin{cases} \partial_t \xi + u \cdot \nabla \xi = 0, \\ \xi|_{t=0} = \xi_0, \end{cases} \tag{1}$$

where $\xi : \mathbb{T}^2 \to \mathbb{R}$ is the vorticity field and

$$u = K * \xi, \quad \operatorname{div} u = 0,$$

is the solenoidal velocity vector field reconstructed from ξ using the Biot–Savart kernel K:

$$K * \xi = -\nabla^{\perp}(-\Delta)^{-1}\xi.$$

Communicated by Lydia Bieri.

✉ Franco Flandoli
franco.flandoli@sns.it

Umberto Pappalettera
umberto.pappalettera@sns.it

1 Scuola Normale Superiore, Piazza dei Cavalieri, 7, 56126 Pisa, Italia

Simulations of this ideal model, as well as observations of roughly two-dimensional physical systems like certain layers of the atmosphere, show a superposition of vortex structures of different size. The basic idea behind this work is that with a great degree of approximation, one could describe the motion of large-scale structures by a stochastic version of 2D Euler equations, where the noise replaces part of the influence of small-scale structures on large-scale ones. This fits with the general idea of stochastic model reduction (Majda et al. 2001; Franzke et al. 2005; Franzke and Majda 2006; Jain et al. 2015; Franzke et al. 2019), but the precise formulation given here is new to our knowledge.

Mathematically speaking, we present a convergence result from a system of two, coupled, Euler-type equations to a single stochastic Euler equation with transport-type Stratonovich noise. Behind the theoretical statement, there is a heuristic motivation based on three time scales, carefully described in Sect. 2.

Let us start with the mathematical result. The system of two, coupled, Euler-type equations we consider in this work is the following:

$$\begin{cases} d\xi_{\mathrm{L}}^{\epsilon} + u_{\mathrm{L}}^{\epsilon} \cdot \nabla \xi_{\mathrm{L}}^{\epsilon} dt = -u_{\mathrm{S}}^{\epsilon} \cdot \nabla \xi_{\mathrm{L}}^{\epsilon} dt, \\ d\xi_{\mathrm{S}}^{\epsilon} + u_{\mathrm{L}}^{\epsilon} \cdot \nabla \xi_{\mathrm{S}}^{\epsilon} dt = -\epsilon^{-2} \xi_{\mathrm{S}}^{\epsilon} dt + \epsilon^{-2} dW, \\ u_{\mathrm{L}}^{\epsilon} = K * \xi_{\mathrm{L}}^{\epsilon}, \\ u_{\mathrm{S}}^{\epsilon} = K * \xi_{\mathrm{S}}^{\epsilon}, \\ \xi_{\mathrm{L}}^{\epsilon}|_{t=0} = \xi_0, \quad \xi_{\mathrm{S}}^{\epsilon}|_{t=0} = \xi_{\mathrm{S}}^{0,\epsilon}. \end{cases} \tag{E}$$

Here, $\epsilon > 0$ is a scaling parameter and $(W_t)_{t \geq 0}$ is a space-dependent Brownian motion of the form:

$$W_t(x) = \sum_{k \in \mathbb{N}} \theta_k(x) \beta_t^k, \tag{2}$$

where the family $\{\beta^k\}_{k \in \mathbb{N}}$ is made of independent standard Brownian motions and the coefficients θ_k are solenoidal, periodic and zero mean, sufficiently regular and decrease sufficiently fast with respect to k, in a suitable sense to be determined later.

The subscripts in the two components $\left(\xi_{\mathrm{L}}^{\epsilon}, \xi_{\mathrm{S}}^{\epsilon}\right)$ refer to large scales and small scales. For the sake of simplicity, we take the initial condition $\xi_{\mathrm{S}}^{0,\epsilon}$ to be distributed as the invariant measure of the linear part of the equation for $\xi_{\mathrm{S}}^{\epsilon}$ (see Sect. 3 for details), but a more general initial condition in the small-scale dynamics can be easily handled.

Our main result is the following, the precise meaning of solution to (E) being given by Proposition 6.1:

Theorem 1.1 *Let $T > 0$ and suppose we are given a zero-mean $\xi_0 \in L^{\infty}(\mathbb{T}^2)$. Denote $B_t(x) = -K * W_t(x)$, and let ξ_L be the unique solution of the stochastic equation*

$$\begin{cases} d\xi_L + u_L \cdot \nabla \xi_L dt = \nabla \xi_L \circ dB, \\ u_L = K * \xi_L, \\ \xi_L|_{t=0} = \xi_0. \end{cases} \tag{3}$$

Then, under suitable assumptions on the coefficients θ_k, the process ξ_L^ϵ solution of (E) *converges as $\epsilon \to 0$ to ξ_L in the following sense: for every $f \in L^1(\mathbb{T}^2)$:*

$$\mathbb{E}\left[\left|\int_{\mathbb{T}^2} \xi_L^\epsilon(t,x) f(x)\,\mathrm{d}x - \int_{\mathbb{T}^2} \xi_L(t,x) f(x)\,\mathrm{d}x\right|\right] \to 0$$

as $\epsilon \to 0$, for every fixed $t \in [0,T]$ and in $L^p([0,T])$ for every finite p. Under the same assumptions on the coefficients θ_k, the velocity field $u_L^\epsilon = K * \xi_L^\epsilon$ converges as $\epsilon \to 0$, in mean value, to $u_L = K * \xi_L$, as variables in $C([0,T], L^1(\mathbb{T}^2, \mathbb{R}^2))$.

Equations of fluid mechanics with Stratonovich transport noise like (3) received great attention in recent years. Precursors already appeared several years ago, see, for instance, (Brzeźniak et al. 1991, 1992; Mikulevicius and Rozovskii 2004, 2005). Then, it was observed, for particular models (see, for instance, Flandoli et al. 2010; Maurelli 2011; Flandoli et al. 2014; Barbato et al. 2014; Bianchi 2013; Beck et al. 2014; Flandoli 2010; Bianchi and Flandoli 2020 and others) that such noise has sometimes rich regularizing properties, typically in terms of improved uniqueness results or blow-up control. This also contributed to additional investigations on such random perturbation. More recently, the problem of which precise Stratonovich transport-advection noise should be considered was understood by Holm (2015) by the development of a stochastic geometric approach based on a variational principle; concerning this important issue, let us mention that the correct noise term for the vorticity equation in three dimensions has the form $\nabla\xi_\mathrm{L} \circ \mathrm{d}B - \xi_\mathrm{L} \circ d\nabla B$, which reduces in 2D to $\nabla\xi_\mathrm{L} \circ \mathrm{d}B$, the noise used in the theorem above (see, for instance, Crisan et al. 2019 for a rigorous result in the 3D case). Our result here, therefore, adds further motivation for the use of this kind of random perturbations; see also (Cotter et al. 2017; Gay-Balmaz and Holm 2018) for a justification of this noise from a viewpoint that has certain conceptual similarities with our one here.

In Sect. 2, we describe in detail why a system for $\left(\xi_\mathrm{L}^\epsilon, \xi_\mathrm{S}^\epsilon\right)$ like (E) above may arise in applications. It is not only a question of splitting the global vorticity field in two parts; a central detail, responsible for the final result, is the precise scaling $\epsilon^{-2}\xi_\mathrm{S}^\epsilon dt + \epsilon^{-2}\mathrm{d}W$. It is not obvious, a priori, why this scaling should appear, since the usual stochastic equations with a scaling parameter that appear in the literature have the form $\epsilon^{-2}\xi_\mathrm{S}^\epsilon \mathrm{d}t + \epsilon^{-1}\mathrm{d}W$. But when there are three time scales in the system, with the features outlined in Sect. 2, the special scaling of our model is natural. See Majda et al. (2001) and subsequent works for similar arguments that were the basis of our research, although other aspects are basically different—in particular the finite dimensionality of the limit models in those works. As remarked in Sect. 2, one issue over others is critical in the approximations: the inverse cascade is not properly captured by this model. This is, however, a general open problem in the realm of stochastic model reduction.

Our final result looks like a particular issue of the general Wong–Zakai approximation principle Wong and Zakai (1965). For the Euler system, it seems the first result in this direction. In the case of Navier–Stokes equations, other forms are already known, see (Hofmanová et al. 2019; Hofmanova et al. 2019) based on rough path theory; for different equations, we mention among others the results contained in Brzeźniak et al. (1988), Twardowska (1993), Tessitore and Zabczyk (2006).

Our proof is based on a probabilistic argument for the Lagrangian dynamics associated with the problem (E): in fact, the formulation itself—the meaning of solution—adopted here is the Lagrangian one. For the deterministic Euler equations, the Lagrangian approach is classical, see, for instance, Marchioro and Pulvirenti (1994) where it is also used to prove existence and uniqueness of a solution of class $\xi \in L^\infty\left([0,T], L^\infty\left(\mathbb{T}^2\right)\right)$ for bounded measurable zero-mean initial vorticity. For the stochastic case, we rely on similar results proved in Brzeźniak et al. (2016).

In the present work, we prove in the first place a convergence result for the Lagrangian particle trajectories, or *characteristics*. Then, relying on the measure-preserving property of characteristics, we are able to prove convergence of the vorticity fields in the sense of Theorem 1.1. We would like to stress the following technical issue: the equation of characteristics contains the velocity field itself as drift, and a careful analysis of the Biot–Savart kernel is required to overcome this difficulty. We hope that our method can be generalized to other equations in dimension two similar to Euler, such as modified surface quasi-geostrophic equations Chae et al. (2011). Three-dimensional models might also be included, possibly requiring a regularization of the nonlinearity as in Cheskidov et al. (2005).

The paper is organized as follows. In Sect. 2, we present the main motivations behind this work; in particular, we justify the interest in the asymptotics as $\epsilon \to 0$ of system (E). In Sect. 3, we introduce a rigorous mathematical setting and give a reformulation of the convergence $\xi_{\mathrm{L}}^\epsilon \to \xi_{\mathrm{L}}$ in terms of the convergence of the characteristics, see below for details; here, we introduce a simplified version of system (E), which is more convenient to capture the main mathematical features of the original system without obscuring them behind heavy calculations. Subsequent Sect. 4 is devoted to the convergence of characteristics, which relies on an argument similar to those contained in Ikeda and Watanabe (1989) as well as some classical estimates on the Biot–Savart kernel K. In Sect. 5, we see how the convergence of the vorticity fields (in the sense of Theorem 1.1) can be deduced from the convergence at the level of characteristics. Finally, in Sect. 6, we transpose the results concerning the simplified system introduced in Sect. 3 back to the original system (E). In the appendix, we prove the equivalence between the Lagrangian notion of solution and the distributional notion of solution to (E), in a sense to be specified later, thus further broadening the scope of our results.

2 Motivations

In this section, we discuss the motivations that justify our interest for the asymptotical behaviour as $\epsilon \to 0$ of $\xi_{\mathrm{L}}^\epsilon$ solution to (E).

First of all, we clarify from the beginning that the theory illustrated in this work applies to systems with *three* time scales, this sentence to be understood as explained below.

We need a small time scale $\mathcal{T}_{\mathrm{S}}$ at which we observe variations, fluctuations, of the main fields (here the vorticity field). We need an intermediate scale $\mathcal{T}_{\mathrm{M}}$ at which the previous fluctuations look random, but not like a white noise, just random with a typical time of variation of order $\mathcal{T}_{\mathrm{S}}$ (small with respect to $\mathcal{T}_{\mathrm{M}}$). Then, we need a third, large, time scale $\mathcal{T}_{\mathrm{L}}$, where, as a result of the theory, the small-scale fluctuations will appear as a white noise, of multiplicative type in the present work. The following relation will play a role:

$$\frac{\mathcal{T}_{\mathrm{L}}}{\mathcal{T}_{\mathrm{M}}} = \frac{\mathcal{T}_{\mathrm{M}}}{\mathcal{T}_{\mathrm{S}}}. \tag{4}$$

We illustrate this framework of three time scales by means of an admittedly phenomenological model. We think of a fluid which develops small-scale fluctuations at the time scale of 1 s: think to wind, roughly two dimensional to fit with our mathematical result, which flows over an irregular ground producing small-scale vortices and perturbations. The small-scale $\mathcal{T}_{\mathrm{S}}$ has the order of 1 s. The intermediate scale has the order of 1 min: in a minute, the fluctuations we observe appear as random, with a typical fluctuation time of 1 s. The large time scale will be of the order of 1 h: at such scale, the fluctuations will look as a white noise.

An example, always ideal, may be the atmospheric fluid over a large region, limited to the lower layer, the one that interacts with the irregularities of the ground (like the mountains). Not aiming to a precise description of such a complex physical system, but just to visualize certain ideas, let us idealize such fluid by means of 2D Euler equations with forcing, written in vorticity form:

$$\begin{cases} \partial_t \xi + u \cdot \nabla \xi = f, \\ u = K * \xi, \end{cases} \tag{5}$$

where f represents the production of small-scale perturbations by the irregularities of hills and mountain profiles, for instance. For long run investigations, it is necessary to include other realistic terms, like a small friction $-\alpha\xi$ and an even smaller dissipation $\nu\Delta\xi$, for some coefficients $1 \gg \alpha \gg \nu > 0$, in order to dissipate the energy introduced by f, but it is not essential to discuss such facts here.

2.1 Human Scale: Seconds

By human scale, we mean the system observed by us, humans, who observe distances in meters and appreciate variations over time spans of seconds. The key quantity here is $\mathcal{T}_{\mathrm{S}} = 1$ s. Velocity $u(t, x)$ is measured in m/s and vorticity $\xi(t, x)$ in s^{-1}.

Assume we split the initial conditions according to some reasonable rule (geometric, spectral...), in large and small scales

$$\xi|_{t=0} = \xi_{\mathrm{L}}(0) + \xi_{\mathrm{S}}(0).$$

Small scales describe the wind fluctuations at space distances of $1 - 10$ m, and large scales are those which impact at the regional level (national, continental), namely with structures of size $10 - 1000$ km. We assume this separation of scales at time $t = 0$.

Having in mind (5), and the previous splitting of the vorticity field in large and small scales, we consider the following system for the evolution of ξ_L, ξ_S:

$$\begin{cases} \partial_t \xi_L + (u_L + u_S) \cdot \nabla \xi_L = 0, \\ \partial_t \xi_S + (u_L + u_S) \cdot \nabla \xi_S = f_S, \end{cases} \tag{6}$$

where $u_L = K * \xi_L$, $u_S = K * \xi_S$ and f_S incorporates the small-scale inputs due to ground irregularities. We assume that f_S includes variations at distances of $1 - 10$ m, with changes in time in a range of order of 1 s.

It is easy to check that the splitting (6) is consistent with (5), in the sense that if (ξ_L, ξ_S) is a solution of (6), then $\xi = \xi_L + \xi_S$ is a solution of (5). We point out, however, that (6) can not be deduced from (5) and the separation of scales at time $t = 0$, but rather it is a modelling hypothesis.

2.2 Intermediate Scale: Minutes

Let us observe the same system from the viewpoint of a recording device which keeps memory of the wind, but with a time scale of minutes: $\mathcal{T}_M = 1$ min. At such time scale, the fluctuations described in the previous subsection look random, the spatial scale being the same as above: $1 - 10$ m.

This motivates our main modelling assumption, see also (Penland and Matrosova 1994; Majda et al. 2001; Boffetta and Ecke 2012). We replace the small scales by a stochastic equation, Gaussian conditionally to the large scales:

$$\begin{cases} \partial_t \xi_L + (u_L + u_S) \cdot \nabla \xi_L = 0, \\ \partial_t \xi_S + u_L \cdot \nabla \xi_S = -\frac{1}{\tau_M} \xi_S + \frac{\sigma}{\sqrt{\tau_M}} W_S', \end{cases} \tag{7}$$

with

$$\tau_M = \frac{1}{60}$$

in the unit of measure of minutes.

Remark 1 We cannot introduce this modelling assumption at the human scale, and it is too unrealistic. If we could, the value of the constant τ would be $\tau_S = 1$, in the unit of measure of seconds.

Heuristically speaking, in order to understand the phenomenology of the second equation, let us drop the term $u_L \cdot \nabla \xi_S$, let us think to W_S as a one dimensional Brownian motion, and realize that the stochastic process defined as

$$\widetilde{\xi}_S(t) := e^{-\frac{1}{\tau_M} t} \xi_S(0) + \int_0^t e^{-\frac{1}{\tau_M}(t-s)} \frac{\sigma}{\sqrt{\tau_M}} W_S'(s),$$

which is a caricature of the true process ξ_S, converges very fast (on the time scale of minutes) to a stationary process, and—similarly—it takes roughly $\tau_M = \frac{1}{60}$ min to go back to equilibrium after a fluctuation. Thus, at the intermediate time scale $\mathcal{T}_{\mathrm{M}}$, the small-scale process looks random, with visible variations every $\frac{1}{60}$ units of time. Its intensity is (essentially) independent of τ_{M} and given by σ: the variance of the stochastic integral in the previous formula is $\int_0^t e^{-\frac{2}{\tau_{\mathrm{M}}}(t-s)} \frac{\sigma^2}{\tau_{\mathrm{M}}} \mathrm{d}s$. When the noise is space-dependent, the intensity is also modulated in space, so σ is a sort of global, mean order of magnitude.

The replacement just discussed of the true small-scale equation by a stochastic equation has some natural motivations, discussed above, but it also has flaws. One of them is related to the inverse cascade, which dominates the energy transfer between scales in 2D, see Kraichnan (1967). Inverse cascade is mostly discarded in this model, having replaced the transfer mechanism due to the term $u_{\mathrm{S}} \cdot \nabla \xi_{\mathrm{S}}$ by a Gaussian term with no Fourier exchange. We do not know how to remedy this drawback. Let us only mention that generally speaking, the problem of a correct energy transfer between scales in stochastic parametrization and stochastic model reduction theories is the most important essentially open problem; our work is not a contribution to the solution of this extremely difficult problem but only the description of a particular stochastic model reduction procedure, different from others previously introduced in the literature.

2.3 Regional Scale: Hours

By this, we mean the same system, lower atmospheric layer over a large region, observed by a satellite. The unit of measure of time is $\mathcal{T}_{\mathrm{L}} = 1$ h, and the unit of measure of space may be $10 - 1000$ km, that is, now different from the spatial scale of meters proper of human and intermediate points of view. We have chosen this scales having in mind, for instance, weather prediction.

How does it look like the system above seen at this space-time scale? If there is no noise term, the formulae are the same as above with

$$\tau_{\mathrm{L}} = \frac{1}{60 \times 60}.$$

But this rescaling, correct for the term $-\frac{1}{\tau_{\mathrm{L}}} \xi_{\mathrm{S}}$, does not hold true for the stochastic term $\frac{\sigma}{\sqrt{\tau}} W_{\mathrm{S}}'$. Let us see more closely the correct rescaling.

Remark 2 We start to see here how the final result depends on the precise procedure described in this section. If we had imposed the stochastic structure of small scales from the very beginning, namely at the human level, then the intermediate step would be unessential, since a single rescaling to the regional level would give the same result. But the result of this alternative procedure would not be the one described in this work, it would be different. It is essential that the passage from human to intermediate scale is based on the rules of deterministic calculus, while the passage from the intermediate to the regional scales is based on the rules of stochastic calculus. Only in this way we get the scaling factors characteristic of the theory described in this work.

In order to avoid trivial mistakes in the rescaling from intermediate to regional scale, let us formalize in more detail the change of unit of measure. The space and time variables at intermediate level will be denoted by x, t and those at regional level by X, T. Essential is that the unit of measure of t is minutes and the one of T is hours, differing by the factor

$$\epsilon^{-1} = 60, \quad t = \epsilon^{-1} T.$$

Less essential here is the role of the unit of measures of x and X. We assume they differ by a factor ϵ_x^{-1}, namely

$$x = \epsilon_x^{-1} X.$$

The only place relevant for applications where it will appear is in the modification of the space-covariance of the noise, which, however, is not our main concern here.

Denote by $u\,(t, x)$ and $U\,(T, X)$ the velocities in the intermediate and regional scale, respectively, and similarly by $\xi\,(t, x)$ and $\Xi\,(T, X)$ for the vorticities. We adopt the same notation for their large-scale components $u_{\mathrm{L}}, U_{\mathrm{L}}, \xi_{\mathrm{L}}, \Xi_{\mathrm{L}}$ and their small-scale components $u_{\mathrm{S}}, U_{\mathrm{S}}, \xi_{\mathrm{S}}, \Xi_{\mathrm{S}}$. We have

$$U\,(T, X) = \epsilon_x \epsilon^{-1} u \left(\frac{T}{\epsilon}, \frac{X}{\epsilon_x} \right)$$

and thus

$$\Xi\,(T, X) = \epsilon^{-1} \xi \left(\frac{T}{\epsilon}, \frac{X}{\epsilon_x} \right).$$

Notice that the material derivative preserves its structure under unit measure change, here up to the factor ϵ^{-2}:

$$\left[\frac{\partial \Xi}{\partial T} + U \cdot \nabla_X \Xi \right] (T, X) = \epsilon^{-2} \left[\frac{\partial \xi}{\partial t} + u \cdot \nabla_x \xi \right] \left(\frac{T}{\epsilon}, \frac{X}{\epsilon_x} \right).$$

Similar identities hold for "mixed" material derivatives, like $U_{\mathrm{S}} \cdot \nabla_X \Xi_{\mathrm{L}}$ and $U_{\mathrm{L}} \cdot \nabla_X \Xi_{\mathrm{S}}$.

Now, let us write Eq. (7) from the viewpoint of the satellite:

$$\begin{aligned} &[\partial_T \Xi_{\mathrm{L}} + (U_{\mathrm{L}} + U_{\mathrm{S}}) \cdot \nabla_X \Xi_{\mathrm{L}}]\,(T, X) = 0 \\ &[\partial_T \Xi_{\mathrm{S}} + U_{\mathrm{L}} \cdot \nabla_X \Xi_{\mathrm{S}}]\,(T, X) = \epsilon^{-2} \left[-\frac{1}{\tau_{\mathrm{M}}} \xi_{\mathrm{S}} + \frac{\sigma}{\sqrt{\tau_{\mathrm{M}}}} W_{\mathrm{S}}' \right] \left(\frac{T}{\epsilon}, \frac{X}{\epsilon_x} \right). \end{aligned}$$

Notice that we still have τ_{M} in these equations. Let us elaborate the term on the right-hand side of the second equation. First,

$$-\epsilon^{-2}\frac{1}{\tau_{\mathrm{M}}}\xi_{\mathrm{S}}\left(\frac{T}{\epsilon},\frac{X}{\epsilon_x}\right) = -\frac{1}{\tau_{\mathrm{M}}\epsilon}\,\Xi_{\mathrm{S}}\left(T,X\right)$$
$$=: -\frac{1}{\tau_{\mathrm{L}}}\,\Xi_{\mathrm{S}}\left(T,X\right)$$

having defined

$$\tau_{\mathrm{L}} = \tau_{\mathrm{M}}\epsilon = \frac{1}{60\times 60}.$$

Second, working with finite increments which is more clear when we deal with Brownian motion, we have

$$\begin{aligned}\frac{\Delta W_{\mathrm{S}}}{\Delta t}\left(\frac{T}{\epsilon},\frac{X}{\epsilon_x}\right) &= \frac{W_{\mathrm{S}}\left(\frac{T}{\epsilon}+\Delta t,\frac{X}{\epsilon_x}\right)-W_{\mathrm{S}}\left(\frac{T}{\epsilon},\frac{X}{\epsilon_x}\right)}{\Delta t}\\ &= \epsilon\,\frac{W_{\mathrm{S}}\left(\frac{T+\epsilon\Delta t}{\epsilon},\frac{X}{\epsilon_x}\right)-W_{\mathrm{S}}\left(\frac{T}{\epsilon},\frac{X}{\epsilon_x}\right)}{\epsilon\Delta t}\\ &\overset{\mathcal{L}}{=} \sqrt{\epsilon}\,\frac{\widetilde{W}_{\mathrm{S}}\left(T+\epsilon\Delta t,\frac{X}{\epsilon_x}\right)-\widetilde{W}_{\mathrm{S}}\left(T,\frac{X}{\epsilon_x}\right)}{\epsilon\Delta t},\end{aligned}$$

for an auxiliary Brownian motion $\widetilde{W}_{\mathrm{S}}$, namely

$$W'_{\mathrm{S}}\left(\frac{T}{\epsilon},\frac{X}{\epsilon_x}\right) \overset{\mathcal{L}}{=} \sqrt{\epsilon}\,\widetilde{W}'_{\mathrm{S}}\left(T,\frac{X}{\epsilon_x}\right),$$

and therefore

$$\epsilon^{-2}\frac{\sigma}{\sqrt{\tau_{\mathrm{M}}}}W'_{\mathrm{S}}\left(\frac{T}{\epsilon},\frac{X}{\epsilon_x}\right) \overset{\mathcal{L}}{=} \epsilon^{-3/2}\frac{\sigma}{\sqrt{\tau_{\mathrm{M}}}}\widetilde{W}'_{\mathrm{S}}\left(T,\frac{X}{\epsilon_x}\right).$$

Hence, the equation for U_{S} reads

$$\left[\partial_T U_{\mathrm{S}} + U_{\mathrm{L}}\cdot\nabla_X U_{\mathrm{S}}\right](T,X) = -\frac{1}{\tau_{\mathrm{L}}}\,\Xi_{\mathrm{S}}\left(T,X\right) + \frac{\sigma}{\epsilon\sqrt{\tau_{\mathrm{L}}}}\widetilde{W}'_{\mathrm{S}}\left(T,\frac{X}{\epsilon_x}\right).$$

The distance at which we still may feel a correlation of the noise $\widetilde{W}'_{\mathrm{S}}\left(T,\frac{X}{\epsilon_x}\right)$ is of the order ϵ_x, rescaled with respect to the intermediate level.

Recall now condition (4). Translated into the new constant it corresponds to what we have tacitly assumed, namely that $\epsilon = \tau_{\mathrm{M}}$. This implies $\tau_{\mathrm{L}} = \epsilon^2$, and thus

$$\left[\partial_T U_{\mathrm{S}} + U_{\mathrm{L}}\cdot\nabla_X U_{\mathrm{S}}\right](T,X) = -\frac{1}{\epsilon^2}\,\Xi_{\mathrm{S}}\left(T,X\right) + \frac{\sigma}{\epsilon^2}\widetilde{W}'_{\mathrm{S}}\left(T,\frac{X}{\epsilon_x}\right),$$

which is the form of our starting model of the rigorous theory (with different notation).

3 Notation and Preliminaries

For any $p \in [1, \infty]$ denote $L^p_0(\mathbb{T}^2)$ the space of p-integrable zero-mean real functions on the two-dimensional torus $\mathbb{T}^2$. For the sake of a clear and effective presentation, we decide to study in the first place the following *simplified* 2D Euler system:

$$\begin{cases} \mathrm{d}\xi^\epsilon_t + u^\epsilon_t \cdot \nabla \xi^\epsilon_t \mathrm{d}t = -\sum_{k\in\mathbb{N}} \sigma_k \cdot \nabla \xi^\epsilon_t \eta^{\epsilon,k}_t \mathrm{d}t, \\ u^\epsilon_t = K * \xi^\epsilon_t, \\ \xi^\epsilon|_{t=0} = \xi_0, \end{cases} \tag{sE}$$

where $\xi_0 \in L^\infty_0(\mathbb{T}^2)$ is the (deterministic) initial condition, K is the Biot–Savart kernel on the two-dimensional torus $\mathbb{T}^2$, $\sigma_k = K * \theta_k : \mathbb{T}^2 \to \mathbb{R}^2$ and $\eta^{\epsilon,k}$ is a Ornstein–Uhlenbeck process:

$$\eta^{\epsilon,k}_t = e^{-\epsilon^{-2}t}\eta^{\epsilon,k}_0 + \int_0^t \epsilon^{-2} e^{-\epsilon^{-2}(t-s)} \mathrm{d}\beta^k_s, \qquad k \in \mathbb{N}.$$

The family $\beta = \{\beta^k\}_{k\in\mathbb{N}}$ is made of independent standard Brownian motions on a given filtered probability space $(\Omega, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$, and the initial conditions $\{\eta^{\epsilon,k}_0\}_{k\in\mathbb{N}}$ are measurable with respect to $\mathcal{F}_0$, so that the processes $\{\eta^{\epsilon,k}\}_{k\in\mathbb{N}}$ are progressively measurable with respect to the filtration $\{\mathcal{F}_t\}_{t\geq 0}$. Moreover, up to a possible enlargement of the filtration $\{\mathcal{F}_t\}_{t\geq 0}$, to simplify our discussion we take independent initial conditions $\{\eta^{\epsilon,k}_0\}_{k\in\mathbb{N}}$, also independent of β, and distributed as centred Gaussian variables with variance equal to $\epsilon^{-2}/2$. In this way, the processes $\{\eta^{\epsilon,k}\}_{k\in\mathbb{N}}$ are stationary, and $\eta^{\epsilon,k}$ is independent of $\eta^{\epsilon,h}$ for $k \neq h$.

Remark 3 Notice that the process $\sum_k \sigma_k \eta^{\epsilon,k}$ is nothing but a rough approximation for the small-scale vorticity ξ^ϵ_S, obtained by simply dropping the nonlinear term in the second equation of (E). This simplified formulation of (E) clarifies *why* we expect a Wong–Zakai result to be true for the large-scale vorticity: indeed, for every $k \in \mathbb{N}$ the process $\eta^{\epsilon,k}$ formally converges to a white-in-time noise, because of the following computation:

$$\begin{aligned} \int_0^t \eta^{\epsilon,k}_s \mathrm{d}s &= \int_0^t e^{-\epsilon^{-2}s}\eta^{\epsilon,k}_0 \mathrm{d}s + \int_0^t \left(\int_0^s \epsilon^{-2} e^{-\epsilon^{-2}(s-r)} \mathrm{d}\beta^k_r \right) \mathrm{d}s \\ &= \int_0^t e^{-\epsilon^{-2}s}\eta^{\epsilon,k}_0 \mathrm{d}s + \int_0^t \left(\int_r^t \epsilon^{-2} e^{-\epsilon^{-2}(s-r)} \mathrm{d}s \right) \mathrm{d}\beta^k_r \\ &= \int_0^t e^{-\epsilon^{-2}s}\eta^{\epsilon,k}_0 \mathrm{d}s + \int_0^t \left(1 - e^{-\epsilon^{-2}(t-r)} \mathrm{d}s \right) \mathrm{d}\beta^k_r \\ &= \beta^k_t + O(\epsilon). \end{aligned}$$

We make the following assumption on the coefficients σ_k:

(A1) $\sigma_k \in C^2(\mathbb{T}^2, \mathbb{R}^2)$ for every $k \in \mathbb{N}$ and $\sum_{k\in\mathbb{N}} \|\nabla^2 \sigma_k\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^8)} < \infty$, where $\nabla^2\sigma_k$ is the second spatial derivative of σ_k and is understood as a vector field taking values in $\mathbb{R}^8$:

$$(\nabla^2\sigma_k(x))^{\alpha}_{\beta,\gamma} = \partial_{x_\beta}\partial_{x_\gamma}\sigma^{\alpha}_k(x), \quad x \in \mathbb{T}^2, \quad \alpha, \beta, \gamma \in \{1, 2\}.$$

Assumption (A1) above is immediately translated in the equivalent assumption on the coefficients θ_k of (2):

(A1) $\theta_k \in L^2_0(\mathbb{T}^2) \cap C^1(\mathbb{T}^2, \mathbb{R})$ for every $k \in \mathbb{N}$ and $\sum_{k\in\mathbb{N}} \|\nabla\theta_k\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^2)} < \infty$.

As an example, one can take, for $\mathbf{k} \in \mathbb{Z}^2 \setminus \{(0, 0)\}$ and $e_{\mathbf{k}}(x) = \exp(2\pi i\mathbf{k} \cdot x)$,

$$\theta_{\mathbf{k}}(x) = q_{\mathbf{k}} e_{\mathbf{k}}(x), \qquad q_{\mathbf{k}} \sim \frac{1}{|\mathbf{k}|^{3+\delta}}, \text{ for some } \delta > 0.$$

In order to study well-posedness of the system (sE), we first need to specify what is the notion of solution we are going to study. We give the following definitions:

Definition 3.1 We say that a measurable map $\varphi : \Omega \times [0, T] \times \mathbb{T}^2 \to \mathbb{T}^2$ is a *stochastic flow of homeomorphisms* if:

- For almost every $\omega \in \Omega$, $\varphi(\omega, t) : \mathbb{T}^2 \to \mathbb{T}^2$ is a homeomorphism for every $t \in [0, T]$;
- For every $x \in \mathbb{T}^2$, $\varphi(x) : \Omega \times [0, T] \to \mathbb{T}^2$ is progressively measurable with respect the filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$.

Definition 3.2 A process $\xi \in L^\infty(\Omega \times [0, T] \times \mathbb{T}^2)$ is said to be *weakly progressively measurable* if for every test function $f \in L^1(\mathbb{T}^2)$ the process

$$t \mapsto \int_{\mathbb{T}^2} \xi_t(x) f(x)\mathrm{d}x$$

is progressively measurable with respect to the filtration $\{\mathcal{F}_t\}_{t\geq 0}$.

The notion of solution to (sE) we adopt hereafter is the following Lagrangian formulation:

Definition 3.3 Let $\epsilon > 0$ and $\xi_0 \in L^\infty(\mathbb{T}^2)$. We say that a weakly progressively measurable process ξ^ϵ is a solution to (sE) if it is given by the transportation of the initial vorticity ξ_0 along the particle trajectories, in formulae:

$$\xi^\epsilon_t = \xi_0 \circ (\varphi^\epsilon_t)^{-1}, \tag{8}$$

where $\varphi^\epsilon_t : \mathbb{T}^2 \to \mathbb{T}^2$ is a stochastic flow of homeomorphisms which satisfies for every $x \in \mathbb{T}^2$:

$$\begin{cases} \mathrm{d}\varphi^\epsilon_t(x) = u^\epsilon_t\left(\varphi^\epsilon_t(x)\right)\mathrm{d}t + \sum_{k\in\mathbb{N}} \sigma_k\left(\varphi^\epsilon_t(x)\right)\eta^{\epsilon,k}_t \mathrm{d}t, \\ \varphi^\epsilon_0(x) = x. \end{cases} \tag{9}$$

We adopt the same terminology, *mutatis mutandis*, for equations and systems similar to (sE).

In fact, we shall prove that in this setting there exists an unique triple $(\xi^\epsilon, u^\epsilon, \varphi^\epsilon)$ such that $u^\epsilon = K * \xi^\epsilon$, (8), and (9) hold simultaneously for every $t \in [0, T]$, see Proposition 3.4.

The maps $\mathbb{T}^2 \ni x \mapsto \varphi_t(x)$, $t \in [0, T]$ are usually called the *characteristics* associated with (sE), since they describe the trajectory of an ideal fluid particle with initial position $x_0 = x$.

Proposition 3.4 *Assume (A1). Then, for every $\epsilon > 0$ and $\xi_0 \in L^\infty(\mathbb{T}^2)$ there exists a unique stochastic flow of homeomorphisms φ^ϵ such that (9) holds with $u^\epsilon_t = K * \xi^\epsilon_t$ and*

$$\xi^\epsilon_t = \xi_0 \circ (\varphi^\epsilon_t)^{-1},$$

as variables in $L^\infty(\mathbb{T}^2)$. In particular, system (sE) is well posed in the sense of Definition 3.3.

Moreover, for a.e. $\omega \in \Omega$, the map $\varphi^\epsilon_t : \mathbb{T}^2 \to \mathbb{T}^2$ is measure-preserving with respect to the Lebesgue measure on $\mathbb{T}^2$ for every $t \in [0, T]$:

$$\int_{\mathbb{T}^2} f(x)\mathrm{d}x = \int_{\mathbb{T}^2} f(\varphi^\epsilon_t(y))\mathrm{d}y, \qquad \text{for every } f \in L^1(\mathbb{T}^2).$$

The proof of the previous proposition is omitted, being easily reconstructed from the proof of the analogous result for the characteristics of the full system (E), see Proposition 6.1. Thanks to this proposition, we can finally define our notion of solution.

The presumed limit equation for ξ^ϵ is

$$\begin{cases} \mathrm{d}\xi_t + u_t \cdot \nabla \xi_t \mathrm{d}t = -\sum_{k\in\mathbb{N}} \sigma_k \cdot \nabla \xi_t \circ \mathrm{d}\beta^k_t, \\ u_t = K * \xi_t, \\ \xi|_{t=0} = \xi_0, \end{cases} \tag{10}$$

where $\circ d\beta^k_t$ stands for the Stratonovich integral. In (Brzeźniak et al. 2016, Section 7), it is proved that for C^2 coefficients σ_k, equation (10) admits an unique weakly progressively measurable solution given by

$$\xi_t = \xi_0 \circ (\varphi_t)^{-1}, \tag{11}$$

as variables in $L^\infty(\mathbb{T}^2)$, where φ_t is the stochastic flow of measure-preserving homeomorphisms solution to the SDE

$$\begin{cases} \mathrm{d}\varphi_t(x) = u_t(\varphi_t(x))\,\mathrm{d}t + \sum_k \sigma_k(\varphi_t(x)) \circ \mathrm{d}\beta^k_t, \\ \varphi_0(x) = x. \end{cases} \tag{12}$$

3.1 Reformulation of the Problem

Recall that, since both φ_t^ϵ and φ_t are measure-preserving maps of the torus $\mathbb{T}^2$, for every test function $f \in L^1(\mathbb{T}^2)$ we have the following identities:

$$\int_{\mathbb{T}^2} \xi_t^\epsilon(x) f(x)\,\mathrm{d}x = \int_{\mathbb{T}^2} \xi_0(y)\, f\left(\varphi_t^\epsilon(y)\right) \mathrm{d}y,$$
$$\int_{\mathbb{T}^2} \xi_t(x) f(x)\,\mathrm{d}x = \int_{\mathbb{T}^2} \xi_0(y)\, f(\varphi_t(y))\, \mathrm{d}y.$$

This motivates, in view of the meaning of convergence $\xi^\epsilon \to \xi$ (in a suitable sense), to investigate instead the convergence of characteristics $\varphi^\epsilon \to \varphi$, where the characteristics $\varphi^\epsilon, \varphi$ are stochastic flows of measure-preserving homeomorphisms that solve:

$$\mathrm{d}\varphi_t^\epsilon(x) = u_t^\epsilon\left(\varphi_t^\epsilon(x)\right) \mathrm{d}t + \sum_{k\in\mathbb{N}} \sigma_k\left(\varphi_t^\epsilon(x)\right) \eta_t^{\epsilon,k}\mathrm{d}t,$$
$$\mathrm{d}\varphi_t(x) = u_t(\varphi_t(x))\,\mathrm{d}t + \sum_{k\in\mathbb{N}} \sigma_k(\varphi_t(x)) \circ \mathrm{d}\beta_t^k,$$

keeping in mind, however, that u^ϵ and u are not given functions, but they depend on the other variables; in particular, they are random. Indeed, we do not know a priori that $u^\epsilon \to u$ in some sense, but this information is part of the problem (cfr. Corollary 4.9).

3.2 Properties of the Biot–Savart Kernel

Here, we briefly recall some useful properties of the Biot–Savart kernel K. We refer to Brzeźniak et al. (2016) for details and proofs.

First of all, recall that the convolution $K * \xi$ is well defined for every $\xi \in L^p(\mathbb{T}^2)$, $p \in [1, \infty]$ and the following estimate holds: for every $p \in [1, \infty]$, there exists a constant C such that for every $\xi \in L^p(\mathbb{T}^2)$

$$\|K * \xi\|_{L^p(\mathbb{T}^2,\mathbb{R}^2)} \le C\|\xi\|_{L^p(\mathbb{T}^2)}.$$

For $p \in (1, \infty)$ and $\xi \in L_0^p(\mathbb{T}^2)$, the convolution with K actually represents the Biot–Savart operator:

$$K * \xi = -\nabla^\perp(-\Delta)^{-1}\xi,$$

which, to every $\xi \in L_0^p(\mathbb{T}^2)$, associates the unique zero-mean, divergence-free velocity vector field $u \in W^{1,p}(\mathbb{T}^2, \mathbb{R}^2)$ such that

$$\mathrm{curl}\, u = \xi.$$

Moreover, for every $p \in (1,\infty)$, there exist constants c, C such that for every $\xi \in L^p_0(\mathbb{T}^2)$

$$c\|\xi\|_{L^p(\mathbb{T}^2)} \le \|K * \xi\|_{W^{1,p}(\mathbb{T}^2,\mathbb{R}^2)} \le C\|\xi\|_{L^p(\mathbb{T}^2)}.$$

Let $r \ge 0$. Denote γ the concave function:

$$\gamma(r) = r(1-\log r)\mathbf{1}_{\{0<r<1/e\}} + (r+1/e)\mathbf{1}_{\{r\ge 1/e\}}.$$

The following two lemmas are proved in Brzeźniak et al. (2016).

Lemma 3.5 *There exists a positive constant C such that:*

$$\int_{\mathbb{T}^2} \left|K(x-y) - K(x'-y)\right| \mathrm{d}y \le C\gamma(|x-x'|)$$

for every $x, x' \in \mathbb{T}^2$.

Lemma 3.6 *Fix $T > 0$ and let $\lambda > 0$, $z_0 \in [0, \exp(1-2e^{\lambda T})]$ be constants. Denote z the unique solution of the following ODE:*

$$z_t = z_0 + \lambda \int_0^t \gamma(z_s)\mathrm{d}s.$$

Then, for every $t \in [0,T]$, the following estimate holds:

$$z_t \le e z_0^{\exp(-\lambda t)}.$$

Hereafter, the symbol $\lesssim$ will be used to indicate an inequality up to a multiplicative constant C which depends only of the data of the problem (*e.g.* T, ξ_0, θ_k etc.). However, for the sake of clarity, we always try to show in the calculations where assumption (A1) comes into play.

4 Convergence of Characteristics

For a given $y \in \mathbb{T}^2$, denote $|y|$ the geodesic distance on the flat two-dimensional torus of the point y from $(0,0) \in \mathbb{T}^2$. To keep the notation as simple as possible, we define, for a measurable map φ from $\mathbb{T}^2$ to itself, the following quantity:

$$\|\varphi\|_{L^1(\mathbb{T}^2,\mathbb{T}^2)} = \int_{\mathbb{T}^2} |\varphi(x)|\,\mathrm{d}x.$$

We adopt this notation because of the similarity with the norm of the Banach space $L^1(\mathbb{T}^2,\mathbb{R}^2)$, although $\|\cdot\|_{L^1(\mathbb{T}^2,\mathbb{T}^2)}$ is not a norm on the space of measurable maps

$\mathbb{T}^2 \to \mathbb{T}^2$; in particular, it is not positively homogeneous. In a similar fashion, we define $\|\cdot\|_{L^\infty(\mathbb{T}^2,\mathbb{T}^2)}$ as

$$\|\varphi\|_{L^\infty(\mathbb{T}^2,\mathbb{T}^2)} = \operatorname{ess\,sup}_{x\in\mathbb{T}^2} |\varphi(x)| .$$

In this section, we prove the following result, concerning convergence of the characteristics φ^ϵ of the simplified system (sE) towards the characteristics φ of system (10).

Proposition 4.1 *Assume (A1). Let φ^ϵ be the solution of* (9)*, and let φ be the solution of* (12)*. Then, for every $T > 0$, the following convergence holds as $\epsilon \to 0$:*

$$\mathbb{E}\left[\sup_{s\le T} \left\|\varphi^\epsilon_s - \varphi_s\right\|_{L^1(\mathbb{T}^2,\mathbb{T}^2)}\right] \to 0. \tag{13}$$

The strategy of the proof is the following and is taken from Ikeda and Watanabe (1989), see also Assing et al. (2020). The idea is to discretize the time interval $[0, T]$ into subintervals of the form $[n\delta_\epsilon, (n+1)\delta_\epsilon]$, for a suitable choice of the mesh δ_ϵ. Then, we adapt an argument in Ikeda and Watanabe (1989) that gives a control of the noisy part of the equations for the characteristics φ^ϵ, φ in the regime $\delta_\epsilon^2/\epsilon^3 \to 0$, $\delta_\epsilon/\epsilon^2 \to \infty$. The nonlinear drift is controlled by Lemma 3.5. Finally, Lemma 3.6 gives the convergence (13).

4.1 Estimates on the Increments

In this paragraph, we give some preliminary estimates on the increments of the characteristics φ^ϵ, φ. We will make use of the following lemma on the supremum of the Ornstein–Uhlenbeck process, which can be found in Jia and Zhao (2020).

Lemma 4.2 *Let $T > 0$, $p \ge 1$. Then, for every $k \in \mathbb{N}$:*

$$\mathbb{E}\left[\sup_{s\in[0,T]} |\eta^{\epsilon,k}_s|^p\right] \lesssim \epsilon^{-p}\log^{p/2}(1+\epsilon^{-2}).$$

Hereafter, we absorb every factor $\log(1+\epsilon^{-2})$ coming from Lemma 4.2 in the symbol $\lesssim$. Since we are only interested in the limit $\epsilon \to 0$, the reader can readily check that doing so does not affect the correctness of our next computations.

The first lemma we prove is the following: it permits to control small-time excursions of the characteristics φ^ϵ, in terms of the time increment Δ and the parameter ϵ.

Lemma 4.3 *Let $T > 0$, $p \ge 1$, $\Delta > 0$. Then,*

$$\mathbb{E}\left[\sup_{\substack{t+\delta\le T\\ \delta\le\Delta}} \|\varphi^\epsilon_{t+\delta} - \varphi^\epsilon_t\|^p_{L^\infty(\mathbb{T}^2,\mathbb{T}^2)}\right] \lesssim \frac{\Delta^p}{\epsilon^p}.$$

Proof The increment $\varphi^\epsilon_{t+\delta}(x) - \varphi^\epsilon_t(x)$ can be written as:

$$\varphi^\epsilon_{t+\delta}(x) - \varphi^\epsilon_t(x) = \int_t^{t+\delta} u^\epsilon_s(\varphi^\epsilon_s(x))\mathrm{d}s + \int_t^{t+\delta} \sum_{k\in\mathbb{N}} \sigma_k(\varphi^\epsilon_s(x))\eta^{\epsilon,k}_s \mathrm{d}s;$$

therefore, since $\|u^\epsilon_s\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^2)} \lesssim \|\xi_0\|_{L^\infty(\mathbb{T}^2)}$, we have

$$\sup_{t+\delta\le T} \|\varphi^\epsilon_{t+\delta} - \varphi^\epsilon_t\|_{L^\infty(\mathbb{T}^2,\mathbb{T}^2)} \lesssim \delta \left(1 + \sum_{k\in\mathbb{N}} \|\sigma_k\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^2)} \sup_{s\in[0,T]} |\eta^{\epsilon,k}_s|\right).$$

The thesis follows by Lemma 4.2. □

The previous lemma can be slightly improved by the following:

Lemma 4.4 *For every $T > 0$, $p \ge 1$ and fixed $n = 0, \dots, T/\delta_\epsilon - 1$, we have*

$$\mathbb{E}\left[\|\varphi^\epsilon_{(n+1)\delta_\epsilon} - \varphi^\epsilon_{n\delta_\epsilon}\|^p_{L^\infty(\mathbb{T}^2,\mathbb{T}^2)}\right] \lesssim \frac{\delta_\epsilon^{2p}}{\epsilon^{2p}} + \delta_\epsilon^{p/2} + \epsilon^p.$$

Proof The increment $\varphi^\epsilon_{(n+1)\delta_\epsilon}(x) - \varphi^\epsilon_{n\delta_\epsilon}(x)$ can be written as:

$$\begin{aligned}\varphi^\epsilon_{(n+1)\delta_\epsilon}(x) - \varphi^\epsilon_{n\delta_\epsilon}(x) &= \int_{n\delta_\epsilon}^{(n+1)\delta_\epsilon} u^\epsilon_s(\varphi^\epsilon_s(x))\mathrm{d}s \\ &\quad + \int_{n\delta_\epsilon}^{(n+1)\delta_\epsilon} \sum_{k\in\mathbb{N}} \left(\sigma_k(\varphi^\epsilon_s(x)) - \sigma_k(\varphi^\epsilon_{n\delta_\epsilon}(x))\right) \eta^{\epsilon,k}_s \mathrm{d}s \\ &\quad + \int_{n\delta_\epsilon}^{(n+1)\delta_\epsilon} \sum_{k\in\mathbb{N}} \sigma_k(\varphi^\epsilon_{n\delta_\epsilon}(x))\eta^{\epsilon,k}_s \mathrm{d}s.\end{aligned}$$

The first term is easy and can be controlled as in Lemma 4.3. The second one is bounded in $L^\infty(\mathbb{T}^2, \mathbb{T}^2)$ uniformly in n by

$$\int_0^{\delta_\epsilon} \sum_{k\in\mathbb{N}} \|\nabla\sigma_k\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^4)} \sup_{t+s\le T} \|\varphi^\epsilon_{t+s} - \varphi^\epsilon_t\|_{L^\infty(\mathbb{T}^2,\mathbb{T}^2)} \sup_{s\in[0,T]} |\eta^{\epsilon,k}_s|\mathrm{d}s,$$

and by Hölder inequality with exponent $q > 1$

$$\begin{aligned}&\mathbb{E}\left[\left(\int_0^{\delta_\epsilon} \sum_{k\in\mathbb{N}} \|\nabla\sigma_k\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^4)} \sup_{t+s\le T} \|\varphi^\epsilon_{t+s} - \varphi^\epsilon_t\|_{L^\infty(\mathbb{T}^2,\mathbb{T}^2)} \sup_{s\in[0,T]} |\eta^{\epsilon,k}_s|\mathrm{d}s\right)^p\right] \\ &\le \delta_\epsilon^{p-1} \left(\sum_{k\in\mathbb{N}} \|\nabla\sigma_k\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^4)}\right)^{p-1} \int_0^{\delta_\epsilon} \sum_{k\in\mathbb{N}} \|\nabla\sigma_k\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^4)}\end{aligned}$$

$$\times \mathbb{E}\left[\sup_{t+s\le T}\|\varphi^\epsilon_{t+s}-\varphi^\epsilon_t\|^{pq}_{L^\infty(\mathbb{T}^2,\mathbb{T}^2)}\right]^{1/q}\mathbb{E}\left[\sup_{s\in[0,T]}|\eta^{\epsilon,k}_s|^{pq'}\right]^{1/q'}\mathrm{d}s$$
$$\lesssim \delta_\epsilon^{p-1}\int_0^{\delta_\epsilon}\frac{s^p}{\epsilon^{2p}}\mathrm{d}s\lesssim\frac{\delta_\epsilon^{2p}}{\epsilon^{2p}}.$$

The third term is bounded in $L^\infty(\mathbb{T}^2,\mathbb{R}^2)$ by

$$\sum_{k\in\mathbb{N}}\|\sigma_k\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^2)}\left|\int_{n\delta_\epsilon}^{(n+1)\delta_\epsilon}\eta^{\epsilon,k}_s\mathrm{d}s\right| = \sum_{k\in\mathbb{N}}\|\sigma_k\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^2)}\left|\beta^{\epsilon,k}_{(n+1)\delta_\epsilon}-\beta^{\epsilon,k}_{n\delta_\epsilon}\right|,$$

where $\beta^{\epsilon,k}_t$ stands for the integrated Ornstein–Uhlenbeck process:

$$\beta^{\epsilon,k}_t=\int_0^t\eta^{\epsilon,k}_s\mathrm{d}s,\quad t\in[0,T],k\in\mathbb{N}.$$

Using

$$\beta^{\epsilon,k}_{(n+1)\delta_\epsilon}-\beta^{\epsilon,k}_{n\delta_\epsilon}=\beta^k_{(n+1)\delta_\epsilon}-\beta^k_{n\delta_\epsilon}-\epsilon^2\left(\eta^{\epsilon,k}_{(n+1)\delta_\epsilon}-\eta^{\epsilon,k}_{n\delta_\epsilon}\right)$$

and Lemma 4.2 we get

$$\mathbb{E}\left[\left(\sum_{k\in\mathbb{N}}\|\sigma_k\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^2)}\left|\beta^{\epsilon,k}_{(n+1)\delta_\epsilon}-\beta^{\epsilon,k}_{n\delta_\epsilon}\right|\right)^p\right]\lesssim\delta_\epsilon^{p/2}+\epsilon^p.$$

□

Next, we move to the analogous estimate for the limiting characteristics φ. Denote by $c:\mathbb{T}^2\to\mathbb{R}^2$, the following Stratonovich corrector:

$$c(x)=\frac{1}{2}\sum_{k\in\mathbb{N}}\nabla\sigma_k(x)\cdot\sigma_k(x),\quad x\in\mathbb{T}^2,$$

which allows to rewrite (12) in the following Itō form:

$$\begin{cases}\mathrm{d}\varphi_t(x)=u_t(\varphi_t(x))\,dt+c(\varphi_t(x))\,\mathrm{d}t+\sum_k\sigma_k(\varphi_t(x))\,\mathrm{d}\beta^k_t,\\ \varphi_0(x)=x.\end{cases}$$

Lemma 4.5 *Let $T>0$, $p\ge1$, $\Delta>0$. Then, for every fixed $n=0,\dots,T/\delta_\epsilon-1$:*

$$\mathbb{E}\left[\sup_{\delta\le\Delta}\|\varphi_{n\delta_\epsilon+\delta}-\varphi_{n\delta_\epsilon}\|^p_{L^\infty(\mathbb{T}^2,\mathbb{T}^2)}\right]\lesssim\Delta^{p/2}.$$

Proof The increment $\varphi_{n\delta_\epsilon+\delta}(x) - \varphi_{n\delta_\epsilon}(x)$ can be written as:

$$\begin{aligned}\varphi_{n\delta_\epsilon+\delta}(x) - \varphi_{n\delta_\epsilon}(x) = &\int_{n\delta_\epsilon}^{n\delta_\epsilon+\delta} u_s(\varphi_s(x))ds + \int_{n\delta_\epsilon}^{n\delta_\epsilon+\delta} c(\varphi_s(x))\mathrm{d}s \\ &+ \int_{n\delta_\epsilon}^{n\delta_\epsilon+\delta} \sum_{k\in\mathbb{N}} \sigma_k(\varphi_s(x))\mathrm{d}\beta_s^k.\end{aligned}$$

The first two terms are easy and can be handled as usual. On the other hand, using Burkholder–Davis–Gundy inequality, for fixed n and k, the last term is controlled by

$$\mathbb{E}\left[\sup_{\delta\le\Delta}\left\|\int_{n\delta_\epsilon}^{n\delta_\epsilon+\delta}\sigma_k(\varphi_s(x))d\beta_s^k\right\|^p_{L^\infty(\mathbb{T}^2,\mathbb{R}^2)}\right] \lesssim \Delta^{p/2}\|\sigma_k\|^p_{L^\infty(\mathbb{T}^2,\mathbb{R}^2)},$$

hence, for fixed n and for $\alpha = 1 - 1/p$, Hölder inequality with exponent p gives

$$\begin{aligned}&\mathbb{E}\left[\sup_{\delta\le\Delta}\left\|\sum_{k\in\mathbb{N}}\int_{n\delta_\epsilon}^{n\delta_\epsilon+\delta}\sigma_k(\varphi_s(x))d\beta_s^k\right\|^p_{L^\infty(\mathbb{T}^2,\mathbb{R}^2)}\right] \\ &= \mathbb{E}\left[\sup_{\delta\le\Delta}\left\|\sum_{k\in\mathbb{N}}\|\sigma_k\|^\alpha_{L^\infty(\mathbb{T}^2,\mathbb{R}^2)}\|\sigma_k\|^{-\alpha}_{L^\infty(\mathbb{T}^2,\mathbb{R}^2)}\int_{n\delta_\epsilon}^{n\delta_\epsilon+\delta}\sigma_k(\varphi_s(x))d\beta_s^k\right\|^p_{L^\infty(\mathbb{T}^2,\mathbb{R}^2)}\right] \\ &\le \left(\sum_{k\in\mathbb{N}}\|\sigma_k\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^2)}\right)^p \Delta^{p/2}.\end{aligned}$$

□

4.2 The Nakao Method

The argument presented in this paragraph is due to Nakao and can be found, for instance, in Ikeda and Watanabe (1989). Roughly speaking, it allows to exploit the discretization of the equation to show the closeness, in a certain sense to be specified, between the Stratonovich corrector and the iterated integral of the Ornstein–Uhlenbeck process.

First, we need some preparation. For any $n = 0, \dots, T/\delta_\epsilon - 1$, consider the following decomposition:

$$\begin{aligned}&\int_{n\delta_\epsilon}^{(n+1)\delta_\epsilon}\sum_{k\in\mathbb{N}}\sigma_k(\varphi_s^\epsilon(x))\eta_s^{\epsilon,k}\mathrm{d}s \\ &\quad = \int_{n\delta_\epsilon}^{(n+1)\delta_\epsilon}\sum_{k\in\mathbb{N}}\left(\sigma_k(\varphi_s^\epsilon(x)) - \sigma_k(\varphi_{n\delta_\epsilon}^\epsilon(x))\right)\eta_s^{\epsilon,k}\mathrm{d}s\end{aligned}$$

$$
\begin{aligned}
&+ \int_{n\delta_\epsilon}^{(n+1)\delta_\epsilon} \sum_{k\in\mathbb{N}} \sigma_k(\varphi^\epsilon_{n\delta_\epsilon}(x))\eta^{\epsilon,k}_s \mathrm{d}s \\
&= \int_{n\delta_\epsilon}^{(n+1)\delta_\epsilon} \sum_{k\in\mathbb{N}} \left(\int_{n\delta_\epsilon}^{s} \nabla\sigma_k(\varphi^\epsilon_r(x)) \cdot u^\epsilon_r(\varphi^\epsilon_r(x))\mathrm{d}r \right) \eta^{\epsilon,k}_s \mathrm{d}s \\
&+ \int_{n\delta_\epsilon}^{(n+1)\delta_\epsilon} \sum_{k,h\in\mathbb{N}} \left(\int_{n\delta_\epsilon}^{s} \nabla\sigma_k(\varphi^\epsilon_r(x)) \cdot \sigma_h(\varphi^\epsilon_r(x))\eta^{\epsilon,h}_r \mathrm{d}r \right) \eta^{\epsilon,k}_s \mathrm{d}s \\
&+ \int_{n\delta_\epsilon}^{(n+1)\delta_\epsilon} \sum_{k\in\mathbb{N}} \sigma_k(\varphi^\epsilon_{n\delta_\epsilon}(x))\mathrm{d}\beta^k_s \\
&- \int_{n\delta_\epsilon}^{(n+1)\delta_\epsilon} \sum_{k\in\mathbb{N}} \sigma_k(\varphi^\epsilon_{n\delta_\epsilon}(x))\epsilon^2 \mathrm{d}\eta^{\epsilon,k}_s \\
&= I^\epsilon_1(n) + I^\epsilon_2(n) + I^\epsilon_3(n) + I^\epsilon_4(n).
\end{aligned}
$$

We further decompose

$$
\begin{aligned}
I^\epsilon_2(n) &= \int_{n\delta_\epsilon}^{(n+1)\delta_\epsilon} \sum_{k,h\in\mathbb{N}} \Bigg(\int_{n\delta_\epsilon}^{s} \Big(\nabla\sigma_k(\varphi^\epsilon_r(x)) \cdot \sigma_h(\varphi^\epsilon_r(x)) \\
&- \nabla\sigma_k(\varphi^\epsilon_{n\delta_\epsilon}(x)) \cdot \sigma_h(\varphi^\epsilon_{n\delta_\epsilon}(x)) \Big) \eta^{\epsilon,h}_r dr \Bigg) \eta^{\epsilon,k}_s ds \\
&+ \int_{n\delta_\epsilon}^{(n+1)\delta_\epsilon} \sum_{k,h\in\mathbb{N}} \Bigg(\int_{n\delta_\epsilon}^{s} \Big(\nabla\sigma_k(\varphi^\epsilon_{n\delta_\epsilon}(x)) \cdot \sigma_h(\varphi^\epsilon_{n\delta_\epsilon}(x)) \\
&- \nabla\sigma_k(\varphi_{n\delta_\epsilon}(x)) \cdot \sigma_h(\varphi_{n\delta_\epsilon}(x)) \Big) \eta^{\epsilon,h}_r \mathrm{d}r \Bigg) \eta^{\epsilon,k}_s \mathrm{d}s \\
&+ \int_{n\delta_\epsilon}^{(n+1)\delta_\epsilon} \sum_{k,h\in\mathbb{N}} \left(\int_{n\delta_\epsilon}^{s} \nabla\sigma_k(\varphi_{n\delta_\epsilon}(x)) \cdot \sigma_h(\varphi_{n\delta_\epsilon}(x))\eta^{\epsilon,h}_r \mathrm{d}r \right) \eta^{\epsilon,k}_s \mathrm{d}s \\
&= I^\epsilon_{2a}(n) + I^\epsilon_{2b}(n) + I^\epsilon_{2c}(n).
\end{aligned}
$$

Regarding the limiting Stratonovich integral, we can rewrite:

$$
\begin{aligned}
\int_{n\delta_\epsilon}^{(n+1)\delta_\epsilon} \sum_{k\in\mathbb{N}} \sigma_k(\varphi_s(x)) \circ d\beta^k_s &= \int_{n\delta_\epsilon}^{(n+1)\delta_\epsilon} \sum_{k\in\mathbb{N}} \big(\sigma_k(\varphi_s(x)) - \sigma_k(\varphi_{n\delta_\epsilon}(x)) \big) \mathrm{d}\beta^k_s \\
&+ \int_{n\delta_\epsilon}^{(n+1)\delta_\epsilon} \sum_{k\in\mathbb{N}} \sigma_k(\varphi_{n\delta_\epsilon}(x)) d\beta^k_s \\
&+ \int_{n\delta_\epsilon}^{(n+1)\delta_\epsilon} \big(c(\varphi_s(x)) - c(\varphi_{n\delta_\epsilon}(x)) \big) ds \\
&+ \int_{n\delta_\epsilon}^{(n+1)\delta_\epsilon} c(\varphi_{n\delta_\epsilon}(x)) ds
\end{aligned}
$$

$$= J_1^\epsilon(n) + J_2^\epsilon(n) + J_3^\epsilon(n) + J_4^\epsilon(n).$$

Lemma 4.6 *The following inequalities hold:*

$$\mathbb{E}\left[\sup_{m=1,\dots,T/\delta_\epsilon}\left\|\sum_{n=0}^{m-1} I_1^\epsilon(n)\right\|_{L^1(\mathbb{T}^2,\mathbb{R}^2)}\right] \lesssim \frac{\delta_\epsilon}{\epsilon};$$

$$\mathbb{E}\left[\sup_{m=1,\dots,T/\delta_\epsilon}\left\|\sum_{n=0}^{m-1} I_{2a}^\epsilon(n)\right\|_{L^1(\mathbb{T}^2,\mathbb{R}^2)}\right] \lesssim \frac{\delta_\epsilon^2}{\epsilon^3};$$

$$\mathbb{E}\left[\sup_{m=1,\dots,T/\delta_\epsilon}\left\|\sum_{n=0}^{m-1} I_4^\epsilon(n)\right\|_{L^1(\mathbb{T}^2,\mathbb{R}^2)}\right] \lesssim \frac{\delta_\epsilon}{\epsilon} + \frac{\epsilon}{\delta_\epsilon^{1/2}} + \frac{\epsilon^2}{\delta_\epsilon} + \epsilon;$$

$$\mathbb{E}\left[\sup_{m=1,\dots,T/\delta_\epsilon}\left\|\sum_{n=0}^{m-1} J_1^\epsilon(n)\right\|_{L^1(\mathbb{T}^2,\mathbb{R}^2)}\right] \lesssim \delta_\epsilon^{1/2};$$

$$\mathbb{E}\left[\sup_{m=1,\dots,T/\delta_\epsilon}\left\|\sum_{n=0}^{m-1} J_3^\epsilon(n)\right\|_{L^1(\mathbb{T}^2,\mathbb{R}^2)}\right] \lesssim \delta_\epsilon^{1/2}.$$

In particular, all the quantities above go to zero as $\epsilon \to 0$*, under the condition* $\delta_\epsilon^2/\epsilon^3 \to 0$, $\delta_\epsilon/\epsilon^2 \to \infty$.

Proof Consider first $I_1^\epsilon(n)$. Using $\|u_r^\epsilon\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^2)} \lesssim \|\xi_0\|_{L^\infty(\mathbb{T}^2)}$ for every $r \in [0,T]$, we get

$$\begin{aligned}
&\mathbb{E}\left[\sup_{m=1,\dots,T/\delta_\epsilon}\left\|\sum_{n=0}^{m-1} I_1^\epsilon(n)\right\|_{L^1(\mathbb{T}^2,\mathbb{R}^2)}\right]\\
&\lesssim \mathbb{E}\left[\sum_{n=0}^{T/\delta_\epsilon-1}\int_{n\delta_\epsilon}^{(n+1)\delta_\epsilon}\sum_{k\in\mathbb{N}}\|\nabla\sigma_k\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^4)}(s-n\delta_\epsilon)|\eta_s^{\epsilon,k}|\mathrm{d}s\right]\\
&\lesssim \sum_{k\in\mathbb{N}}\|\nabla\sigma_k\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^4)}\delta_\epsilon\mathbb{E}\left[\sup_{s\in[0,T]}|\eta_s^{\epsilon,k}|\right] \lesssim \frac{\delta_\epsilon}{\epsilon}.
\end{aligned}$$

For the term $I_{2a}^\epsilon(n)$, Lemma 4.3 gives:

$$\begin{aligned}
&\mathbb{E}\left[\sup_{m=1,\dots,T/\delta_\epsilon}\left\|\sum_{n=0}^{m-1} I_{2a}^\epsilon(n)\right\|_{L^1(\mathbb{T}^2,\mathbb{R}^2)}\right]\\
&\le \sum_{k,h\in\mathbb{N}}\left(\|\nabla^2\sigma_k\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^8)}\|\sigma_h\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^2)} + \|\nabla\sigma_k\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^4)}\|\nabla\sigma_h\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^4)}\right)
\end{aligned}$$

$$
\begin{aligned}
&\times \mathbb{E}\left[\sum_{n=0}^{T/\delta_\epsilon-1}\int_{n\delta_\epsilon}^{(n+1)\delta_\epsilon}\int_{n\delta_\epsilon}^{s}\|\varphi_r^\epsilon-\varphi_{n\delta_\epsilon}^\epsilon\|_{L^\infty(\mathbb{T}^2,\mathbb{T}^2)}|\eta_r^{\epsilon,h}||\eta_s^{\epsilon,k}|drds\right]\\
&\lesssim \sum_{k,h\in\mathbb{N}}\left(\|\nabla^2\sigma_k\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^8)}\|\sigma_h\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^2)}+\|\nabla\sigma_k\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^4)}\|\nabla\sigma_h\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^4)}\right)\\
&\times\sum_{n=0}^{T/\delta_\epsilon-1}\int_{n\delta_\epsilon}^{(n+1)\delta_\epsilon}\int_{n\delta_\epsilon}^{s}\frac{r-n\delta_\epsilon}{\epsilon^3}drds\lesssim\frac{\delta_\epsilon^2}{\epsilon^3}.
\end{aligned}
$$

The term $I_4^\epsilon(n)$ is treated after a discrete integration by parts, in order to have a better control of the time increment: indeed, Lemma 4.4 gives

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{m=1,\dots,T/\delta_\epsilon}\left\|\sum_{n=0}^{m-1}I_4^\epsilon(n)\right\|_{L^1(\mathbb{T}^2,\mathbb{R}^2)}\right]\\
&\le\mathbb{E}\left[\sup_{m=1,\dots,T/\delta_\epsilon}\left\|\sum_{n=0}^{m-1}\sum_{k\in\mathbb{N}}\sigma_k(\varphi_{n\delta_\epsilon}^\epsilon(x))\epsilon^2\left(\eta_{(n+1)\delta_\epsilon}^{\epsilon,k}-\eta_{n\delta_\epsilon}^{\epsilon,k}\right)\right\|_{L^1(\mathbb{T}^2,\mathbb{R}^2)}\right]\\
&\lesssim\mathbb{E}\left[\sup_{m=1,\dots,T/\delta_\epsilon}\left\|\sum_{n=1}^{m}\sum_{k\in\mathbb{N}}\left(\sigma_k(\varphi_{n\delta_\epsilon}^\epsilon(x))-\sigma_k(\varphi_{(n-1)\delta_\epsilon}^\epsilon(x))\right)\epsilon^2\eta_{n\delta_\epsilon}^{\epsilon,k}\right\|_{L^1(\mathbb{T}^2,\mathbb{R}^2)}\right]\\
&\quad+\mathbb{E}\left[\left\|\sum_{k\in\mathbb{N}}\sigma_k(\varphi_0^\epsilon(x))\epsilon^2\eta_0^{\epsilon,k}\right\|_{L^1(\mathbb{T}^2,\mathbb{R}^2)}\right]\\
&\quad+\mathbb{E}\left[\sup_{m=1,\dots,T/\delta_\epsilon}\left\|\sum_{k\in\mathbb{N}}\sigma_k(\varphi_m^\epsilon(x))\epsilon^2\eta_m^{\epsilon,k}\right\|_{L^1(\mathbb{T}^2,\mathbb{R}^2)}\right]\\
&\lesssim\mathbb{E}\left[\sum_{n=1}^{T/\delta_\epsilon}\sum_{k\in\mathbb{N}}\|\nabla\sigma_k\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^4)}\left\|\varphi_{n\delta_\epsilon}^\epsilon-\varphi_{(n-1)\delta_\epsilon}^\epsilon\right\|_{L^\infty(\mathbb{T}^2,\mathbb{T}^2)}\epsilon^2|\eta_{n\delta_\epsilon}^{\epsilon,k}|\right]+\epsilon\\
&\lesssim\sum_{n=1}^{T/\delta_\epsilon}\sum_{k\in\mathbb{N}}\|\nabla\sigma_k\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^4)}\left(\frac{\delta_\epsilon^2}{\epsilon^2}+\delta_\epsilon^{1/2}+\epsilon\right)\epsilon+\epsilon\\
&\lesssim\frac{\delta_\epsilon}{\epsilon}+\frac{\epsilon}{\delta_\epsilon^{1/2}}+\frac{\epsilon^2}{\delta_\epsilon}+\epsilon.
\end{aligned}
$$

For the remaining terms $J_1^\epsilon(n)$ and $J_3^\epsilon(n)$, we have by Lemma 4.5

$$
\mathbb{E}\left[\sup_{m=1,\dots,T/\delta_\epsilon}\left\|\sum_{n=0}^{m-1}J_1^\epsilon(n)\right\|_{L^1(\mathbb{T}^2,\mathbb{R}^2)}\right]
$$

$$\lesssim \sum_{k\in\mathbb{N}} \|\nabla\sigma_k\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^4)} \mathbb{E}\left[\left(\sum_{n=0}^{T/\delta_\epsilon-1}\int_{n\delta_\epsilon}^{(n+1)\delta_\epsilon} \|\varphi_s-\varphi_{n\delta_\epsilon}\|^2_{L^\infty(\mathbb{T}^2,\mathbb{T}^2)} ds\right)^{1/2}\right]$$
$$\lesssim \sum_{k\in\mathbb{N}} \|\nabla\sigma_k\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^4)} \mathbb{E}\left[\sum_{n=0}^{T/\delta_\epsilon-1}\int_{n\delta_\epsilon}^{(n+1)\delta_\epsilon} \|\varphi_s-\varphi_{n\delta_\epsilon}\|^2_{L^\infty(\mathbb{T}^2,\mathbb{T}^2)} ds\right]^{1/2} \lesssim \delta_\epsilon^{1/2},$$

and similarly

$$\mathbb{E}\left[\sup_{m=1,\dots,T/\delta_\epsilon}\left\|\sum_{n=0}^{m-1} J_3^\epsilon(n)\right\|_{L^1(\mathbb{T}^2,\mathbb{R}^2)}\right]$$
$$\le \sum_{k\in\mathbb{N}}\left(\|\nabla^2\sigma_k\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^8)}\|\sigma_k\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^2)} + \|\nabla\sigma_k\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^4)}\|\nabla\sigma_k\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^4)}\right)$$
$$\times\, \mathbb{E}\left[\sum_{n=0}^{T/\delta_\epsilon-1}\int_{n\delta_\epsilon}^{(n+1)\delta_\epsilon} \|\varphi_s-\varphi_{n\delta_\epsilon}\|_{L^\infty(\mathbb{T}^2,\mathbb{T}^2)} \mathrm{d}s\right] \lesssim \delta_\epsilon^{1/2}.$$

□

Lemma 4.7 *(Nakao) The following inequality holds:*

$$\mathbb{E}\left[\sup_{m=1,\dots,T/\delta_\epsilon}\left\|\sum_{n=0}^{m-1} I_{2c}^\epsilon(n) - J_4^\epsilon(n)\right\|_{L^1(\mathbb{T}^2,\mathbb{R}^2)}\right] \lesssim \frac{\delta_\epsilon}{\epsilon} + \delta_\epsilon^{1/2} + \frac{\epsilon^2}{\delta_\epsilon}.$$

Proof By the very definition of $I_{2c}^\epsilon(n)$, $J_4^\epsilon(n)$, one has

$$I_{2c}^\epsilon(n) = \sum_{k,h\in\mathbb{N}} \nabla\sigma_k(\varphi_{n\delta_\epsilon}(x))\cdot\sigma_h(\varphi_{n\delta_\epsilon}(x))\int_{n\delta_\epsilon}^{(n+1)\delta_\epsilon}\left(\int_{n\delta_\epsilon}^{s}\eta_r^{\epsilon,h}\mathrm{d}r\right)\eta_s^{\epsilon,k}\mathrm{d}s,$$
$$J_4^\epsilon(n) = \sum_{k\in\mathbb{N}} \nabla\sigma_k(\varphi_{n\delta_\epsilon}(x))\cdot\sigma_k(\varphi_{n\delta_\epsilon}(x))\,\frac{\delta_\epsilon}{2}.$$

Therefore, one can decompose the quantity under investigation as follows:

$$\sum_{n=0}^{m-1} I_{2c}^\epsilon(n) - J_4^\epsilon(n)$$
$$= \sum_{n=0}^{m-1}\sum_{k,h\in\mathbb{N}} \nabla\sigma_k(\varphi_{n\delta_\epsilon}(x))\cdot\sigma_h(\varphi_{n\delta_\epsilon}(x))\left(c_{h,k}^n(\delta_\epsilon,\epsilon) - \mathbb{E}\left[c_{h,k}^n(\delta_\epsilon,\epsilon)\mid\mathcal{F}_{n\delta_\epsilon}\right]\right)$$
$$+ \sum_{n=0}^{m-1}\sum_{k,h\in\mathbb{N}} \nabla\sigma_k(\varphi_{n\delta_\epsilon}(x))\cdot\sigma_h(\varphi_{n\delta_\epsilon}(x))\left(\mathbb{E}\left[c_{h,k}^n(\delta_\epsilon,\epsilon)\mid\mathcal{F}_{n\delta_\epsilon}\right] - \frac{\delta_{h,k}}{2}\delta_\epsilon\right), \tag{14}$$

where $\delta_{h,k}$ is the Kronecker delta function and

$$c^n_{h,k}(\delta_\epsilon, \epsilon) = \int_{n\delta_\epsilon}^{(n+1)\delta_\epsilon} \left(\int_{n\delta_\epsilon}^{s} \eta_r^{\epsilon,h} \mathrm{d}r \right) \eta_s^{\epsilon,k} \mathrm{d}s.$$

Notice that $c^n_{h,k}(\delta_\epsilon, \epsilon)$ is measurable with respect to $\mathcal{F}_{(n+1)\delta_\epsilon}$ and has conditional expectation

$$\begin{aligned}
&\mathbb{E}\left[c^n_{h,k}(\delta_\epsilon, \epsilon) \mid \mathcal{F}_{n\delta_\epsilon}\right] \\
&= \int_{n\delta_\epsilon}^{(n+1)\delta_\epsilon} \left(\int_{n\delta_\epsilon}^{s} \mathbb{E}\left[\eta_r^{\epsilon,h} \eta_s^{\epsilon,k} \mid \mathcal{F}_{n\delta_\epsilon}\right] \mathrm{d}r \right) \eta_s^{\epsilon,k} \mathrm{d}s \\
&= \eta_{n\delta_\epsilon}^{\epsilon,h} \eta_{n\delta_\epsilon}^{\epsilon,k} \int_{n\delta_\epsilon}^{(n+1)\delta_\epsilon} \left(\int_{n\delta_\epsilon}^{s} e^{-\epsilon^{-2}(r+s-2n\delta_\epsilon)} \mathrm{d}r \right) \mathrm{d}s \\
&\quad + \delta_{h,k} \int_{n\delta_\epsilon}^{(n+1)\delta_\epsilon} \left(\int_{n\delta_\epsilon}^{s} \frac{\varepsilon^{-2}}{2} \left(e^{-\varepsilon^{-2}(s-r)} - e^{-\varepsilon^{-2}(r+s-2n\delta_\epsilon)} \right) dr \right) \mathrm{d}s,
\end{aligned}$$

where we have used the mild formulation of η^ϵ:

$$\begin{aligned}
\eta_r^{\epsilon,h} &= e^{-\epsilon^{-2}(r-n\delta_\epsilon)} \eta_{n\delta_\epsilon}^{\epsilon,h} + \int_{n\delta_\epsilon}^{r} \epsilon^{-2} e^{-\epsilon^{-2}(r-r')} \mathrm{d}\beta^h_{r'}, \\
\eta_s^{\epsilon,k} &= e^{-\epsilon^{-2}(s-n\delta_\epsilon)} \eta_{n\delta_\epsilon}^{\epsilon,k} + \int_{n\delta_\epsilon}^{s} \epsilon^{-2} e^{-\epsilon^{-2}(s-s')} \mathrm{d}\beta^k_{s'}.
\end{aligned}$$

An elementary computation gives:

$$\begin{aligned}
\mathbb{E}\left[c^n_{h,k}(\delta_\epsilon, \epsilon) \mid \mathcal{F}_{n\Delta}\right] &= \frac{\varepsilon^4}{2} \eta_{n\delta_\epsilon}^{\epsilon,h} \eta_{n\delta_\epsilon}^{\epsilon,k} \left(e^{-\epsilon^{-2}\delta_\epsilon} - 1 \right)^2 \\
&\quad + \frac{\delta_{h,k}}{2} \left(\delta_\epsilon + \epsilon^2 \left(-\frac{3}{2} + 2e^{-\epsilon^{-2}\delta_\epsilon} - \frac{1}{2} e^{-2\epsilon^{-2}\delta_\epsilon} \right) \right). \qquad (15)
\end{aligned}$$

Since the quantity

$$M_m(x) = \sum_{n=0}^{m-1} \sum_{k,h \in \mathbb{N}} \nabla \sigma_k(\varphi_{n\delta_\epsilon}(x)) \cdot \sigma_h(\varphi_{n\delta_\epsilon}(x)) \left(c^n_{h,k}(\delta_\epsilon, \epsilon) - \mathbb{E}\left[c^n_{h,k}(\delta_\epsilon, \epsilon) \mid \mathcal{F}_{n\delta_\epsilon}\right] \right)$$

is a $L^2(\mathbb{T}^2, \mathbb{R}^2)$-valued martingale with respect to the filtration $(\mathcal{F}_{n\delta_\epsilon})_{n \in \mathbb{N}}$ (crf. Pisier 2016), by Doob maximal inequality and martingale property we have the following:

$$\begin{aligned}
\mathbb{E}\left[\sup_{m=1,\dots,T/\delta_\epsilon} \|M_m\|^2_{L^2(\mathbb{T}^2,\mathbb{R}^2)} \right] &\lesssim \mathbb{E}\left[\|M_{T/\delta_\epsilon}\|^2_{L^2(\mathbb{T}^2,\mathbb{R}^2)} \right] \\
&\lesssim \mathbb{E}\left[\sum_{n=0}^{T/\delta_\epsilon - 1} \|M_{m+1} - M_m\|^2_{L^2(\mathbb{T}^2,\mathbb{R}^2)} \right].
\end{aligned}$$

The conditional expectation is a $L^2(\Omega)$-projection, thus for every $n, h, k \in \mathbb{N}$

$$\mathbb{E}\left[\left|c^n_{h,k}(\delta_\epsilon,\epsilon) - \mathbb{E}\left[c^n_{h,k}(\delta_\epsilon,\epsilon) \mid \mathcal{F}_{n\delta_\epsilon}\right]\right|^2\right] \lesssim \mathbb{E}\left[\left|c^n_{h,k}(\delta_\epsilon,\epsilon)\right|^2\right],$$

and therefore

$$\begin{aligned}
&\mathbb{E}\left[\sum_{n=0}^{T/\delta_\epsilon-1} \|M_{m+1}-M_m\|^2_{L^2(\mathbb{T}^2,\mathbb{R}^2)}\right] \\
&\lesssim \sum_{n=0}^{T/\delta_\epsilon-1}\left(\sum_{h,k\in\mathbb{N}} \|\nabla\sigma_k\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^4)}\|\|\sigma_h\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^2)}\right) \\
&\quad\times\left(\sum_{h,k\in\mathbb{N}} \|\nabla\sigma_k\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^4)}\|\|\sigma_h\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^2)}\mathbb{E}\left[\left|c^n_{h,k}(\delta_\epsilon,\epsilon)\right|^2\right]\right) \lesssim \frac{\delta_\epsilon^2}{\epsilon^2}+\delta_\epsilon.
\end{aligned}$$

Moreover, the process

$$N_m(x) = \sum_{n=0}^{m-1}\sum_{k,h\in\mathbb{N}} \nabla\sigma_k(\varphi_{n\delta_\epsilon}(x))\cdot\sigma_h(\varphi_{n\delta_\epsilon}(x))\left(\mathbb{E}\left[c^n_{h,k}(\delta_\epsilon,\epsilon)\mid\mathcal{F}_{n\delta_\epsilon}\right]-\frac{\delta_{h,k}}{2}\delta_\epsilon\right)$$

satisfies

$$\mathbb{E}\left[\sup_{m=1,\dots,T/\delta_\epsilon}\|N_m\|^2_{L^2(\mathbb{T}^2,\mathbb{R}^2)}\right] \lesssim \frac{\epsilon^4}{\delta_\epsilon^2},$$

which is an easy consequence of (15). By (14) and Hölder inequality, we get

$$\begin{aligned}
&\mathbb{E}\left[\sup_{m=1,\dots,T/\delta_\epsilon}\left\|\sum_{n=0}^{m-1} I^\epsilon_{2c}(n)-J^\epsilon_4(n)\right\|_{L^1(\mathbb{T}^2,\mathbb{R}^2)}\right] \\
&\lesssim \mathbb{E}\left[\sup_{m=1,\dots,T/\delta_\epsilon}\|M_m\|_{L^1(\mathbb{T}^2,\mathbb{R}^2)}\right]+\mathbb{E}\left[\sup_{m=1,\dots,T/\delta_\epsilon}\|N_m\|_{L^1(\mathbb{T}^2,\mathbb{R}^2)}\right] \\
&\lesssim \mathbb{E}\left[\sup_{m=1,\dots,T/\delta_\epsilon}\|M_m\|^2_{L^2(\mathbb{T}^2,\mathbb{R}^2)}\right]^{1/2}+\mathbb{E}\left[\sup_{m=1,\dots,T/\delta_\epsilon}\|N_m\|^2_{L^2(\mathbb{T}^2,\mathbb{R}^2)}\right]^{1/2} \\
&\lesssim \frac{\delta_\epsilon}{\epsilon}+\delta_\epsilon^{1/2}+\frac{\epsilon^2}{\delta_\epsilon}.
\end{aligned}$$

□

We conclude this paragraph with the following result.

Lemma 4.8 *The following estimates hold:*

$$\mathbb{E}\left[\sup_{m=1,\dots,N}\left\|\sum_{n=0}^{m-1} I_3^\epsilon(n) - J_2^\epsilon(n)\right\|_{L^1(\mathbb{T}^2,\mathbb{R}^2)}\right] \lesssim \sum_{m=1}^{N} \delta_\epsilon \mathbb{E}\left[\sup_{n=1,\dots,m}\left\|\varphi^\epsilon_{n\delta_\epsilon} - \varphi_{n\delta_\epsilon}\right\|_{L^1(\mathbb{T}^2,\mathbb{T}^2)}\right];$$

$$\mathbb{E}\left[\sup_{m=1,\dots,N}\left\|\sum_{n=0}^{m-1} I_{2b}^\epsilon(n)\right\|_{L^1(\mathbb{T}^2,\mathbb{R}^2)}\right] \lesssim \sum_{m=1}^{N} \delta_\epsilon \mathbb{E}\left[\sup_{n=1,\dots,m}\left\|\varphi^\epsilon_{n\delta_\epsilon} - \varphi_{n\delta_\epsilon}\right\|_{L^1(\mathbb{T}^2,\mathbb{T}^2)}\right].$$

Proof The first estimate is an easy consequence of Burkholder–Davis–Gundy inequality. For the second estimate, one can argue as in Lemma 4.7 to replace the quantity $I_{2b}^\epsilon(n)$ with:

$$\int_{n\delta_\epsilon}^{(n+1)\delta_\epsilon} \left(c(\varphi^\epsilon_{n\delta_\epsilon}) - c(\varphi_{n\delta_\epsilon})\right) \mathrm{d}s = \delta_\epsilon \left(c(\varphi^\epsilon_{n\delta_\epsilon}) - c(\varphi_{n\delta_\epsilon})\right),$$

up to a correction that is infinitesimal as $\epsilon \to 0$. For the latter quantity, the desired inequality is immediate. □

4.3 Proof of Proposition 4.1

We are ready to prove the main result of this section. Recall

$$d\varphi_t^\epsilon(x) = u_t^\epsilon\left(\varphi_t^\epsilon(x)\right)\mathrm{d}t + \sum_{k\in\mathbb{N}} \sigma_k\left(\varphi_t^\epsilon(x)\right)\eta_t^{\epsilon,k}\mathrm{d}t.$$

Since $\varphi^\epsilon : \mathbb{T}^2 \to \mathbb{T}^2$ is measure-preserving, for Lebesgue a.e. $x \in \mathbb{T}^2$:

$$u_t^\epsilon(\varphi_t^\epsilon(x)) = \int_{\mathbb{T}^2} K(\varphi_t^\epsilon(x) - \varphi_t^\epsilon(y))\xi_0(y)\mathrm{d}y,$$

and therefore, we have the following integral formulation for (9)

$$\varphi_t^\epsilon(x) = x + \int_0^t \left(\int_{\mathbb{T}^2} K(\varphi_s^\epsilon(x) - \varphi_s^\epsilon(y))\xi_0(y)\mathrm{d}y\right)\mathrm{d}s + \int_0^t \sum_{k\in\mathbb{N}} \sigma_k(\varphi_s^\epsilon(x))\eta_s^{\epsilon,k}\mathrm{d}s,$$

and similarly for (12)

$$\varphi_t(x) = x + \int_0^t \left(\int_{\mathbb{T}^2} K(\varphi_s(x) - \varphi_s(y))\xi_0(y)dy \right) \mathrm{d}s + \int_0^t \sum_{k\in\mathbb{N}} \sigma_k(\varphi_s(x)) \circ \mathrm{d}\beta_s^k.$$

Proof of Proposition 4.1 For the difference $Z_t^\epsilon(x) = \varphi_t^\epsilon(x) - \varphi_t(x)$, we have:

$$\begin{aligned} Z_t^\epsilon(x) = &\int_0^t \left(\int_{\mathbb{T}^2} \left(K(\varphi_s^\epsilon(x) - \varphi_s^\epsilon(y)) - K(\varphi_s^\epsilon(x) - \varphi_s(y))\right) \xi_0(y)\mathrm{d}y \right) \mathrm{d}s \\ &+ \int_0^t \left(\int_{\mathbb{T}^2} \left(K(\varphi_s^\epsilon(x) - \varphi_s(y)) - K(\varphi_s(x) - \varphi_s(y))\right) \xi_0(y)dy \right) ds \\ &+ \int_0^t \sum_{k\in\mathbb{N}} \sigma_k(\varphi_s^\epsilon(x))\eta_s^{\epsilon,k} ds - \int_0^t \sum_{k\in\mathbb{N}} \sigma_k(\varphi_s(x)) \circ d\beta_s^k. \end{aligned}$$

Using the estimates given by Lemma 4.3 and Lemma 4.5, we can approximate the latter two integrals in the expression above with their discretized versions, computed in a point $n\delta_\epsilon$ such that $t \in [n\delta_\epsilon, (n+1)\delta_\epsilon)$, up to a correction that is infinitesimal as $\epsilon \to 0$. Then, in the regime $\delta_\epsilon^2/\epsilon^3 \to 0$, $\delta_\epsilon/\epsilon^2 \to \infty$, using the results of subsection 4.2, Lemma 3.5 and the concavity of γ, we arrive to

$$\begin{aligned} \mathbb{E}\left[\sup_{s\le t} \|Z_s^\epsilon\|_{L^1(\mathbb{T}^2,\mathbb{T}^2)}\right] &\lesssim \int_0^t \gamma\left(\mathbb{E}\left[\sup_{r\le s} \|Z_r^\epsilon\|_{L^1(\mathbb{T}^2,\mathbb{T}^2)}\right]\right) \mathrm{d}s \\ &\quad + r_T^\epsilon + \sum_{n=1}^{\lfloor t/\delta_\epsilon \rfloor} \delta_\epsilon \mathbb{E}\left[\sup_{r\le n} \|Z_r^\epsilon\|_{L^1(\mathbb{T}^2,\mathbb{T}^2)}\right] \\ &\lesssim \int_0^t \gamma\left(\mathbb{E}\left[\sup_{r\le s} \|Z_r^\epsilon\|_{L^1(\mathbb{T}^2,\mathbb{T}^2)}\right]\right) \mathrm{d}s + r_T^\epsilon, \end{aligned}$$

where r_T^ϵ is a remainder coming from the discretization procedure, Lemma 4.6 and Lemma 4.7, and it goes to zero as $\epsilon \to 0$. By Lemma 3.6, we conclude that

$$\mathbb{E}\left[\sup_{s\le T} \|Z_s^\epsilon\|_{L^1(\mathbb{T}^2,\mathbb{T}^2)}\right] \to 0,$$

and therefore $Z^\epsilon \to 0$ in mean value as a variable in $C([0,T], L^1(\mathbb{T}^2,\mathbb{T}^2))$. □

Lemma 3.5 and the same calculations as above yield the following convergence at the velocity level:

Corollary 4.9 *Assume (A1), and let $u^\epsilon = K * \xi^\epsilon$ (resp. $u = K * \xi$) be the velocity field associated with the characteristics φ^ϵ (resp. φ). Then, as $\epsilon \to 0$:*

$$\mathbb{E}\left[\sup_{s\le T} \|u_s^\epsilon - u_s\|_{L^1(\mathbb{T}^2,\mathbb{R}^2)}\right] \lesssim \gamma\left(\mathbb{E}\left[\sup_{s\le T} \|\varphi_s^\epsilon - \varphi_s\|_{L^1(\mathbb{T}^2,\mathbb{T}^2)}\right]\right) \to 0.$$

5 Convergence of the Vorticity Process

In this brief section, we discuss the consequences of the convergence of the characteristics at the level of the vorticity process. Recall that we have proved:

$$\mathbb{E}\left[\sup_{s\leq T}\left\|\varphi_s^\epsilon-\varphi_s\right\|_{L^1(\mathbb{T}^2,\mathbb{T}^2)}\right]\to 0,$$

as $\epsilon\to 0$. We have the following:

Theorem 5.1 *The vorticity process ξ^ϵ solution of the simplified system (sE) converges to ξ solution of (10) as $\epsilon\to 0$ in the following sense: for every $f\in L^1(\mathbb{T}^2)$:*

$$\mathbb{E}\left[\left|\int_{\mathbb{T}^2}\xi_t^\epsilon(x)f(x)\,dx-\int_{\mathbb{T}^2}\xi_t(x)f(x)\,dx\right|\right]\to 0$$

as $\epsilon\to 0$, for every fixed $t\in[0,T]$ and in $L^p([0,T])$ for every finite p.

Proof By (8) and the fact that $\varphi^\epsilon:\mathbb{T}^2\to\mathbb{T}^2$ is measure-preserving, a change of variable leads to

$$\int_{\mathbb{T}^2}\xi_t^\epsilon(x)f(x)\,\mathrm{d}x=\int_{\mathbb{T}^2}\xi_0(y)\,f\left(\varphi_t^\epsilon(y)\right)\mathrm{d}y,$$

for every $t\in[0,T]$ and $f\in L^1(\mathbb{T}^2)$. Similarly,

$$\int_{\mathbb{T}^2}\xi_t(x)f(x)\,\mathrm{d}x=\int_{\mathbb{T}^2}\xi_0(y)\,f(\varphi_t(y))\,\mathrm{d}y.$$

Since $f\in L^1(\mathbb{T}^2)$, then by Lusin theorem (Rudin 1970, Theorem 2.23) for every $\delta>0$, there exists a continuous function $f_\delta\in C(\mathbb{T}^2)$ and a compact set C_δ such that f coincides with f_δ on C_δ and $meas(\mathbb{T}^2\setminus C_\delta)<\delta$. Therefore,

$$\begin{aligned}&\left|\int_{\mathbb{T}^2}\xi_0(y)\,f\left(\varphi_t^\epsilon(y)\right)\mathrm{d}y-\int_{\mathbb{T}^2}\xi_0(y)\,f(\varphi_t(y))\,\mathrm{d}y\right|\\&\leq\|\xi_0\|_{L^\infty(\mathbb{T}^2)}\int_{C_\delta}\left|f(\varphi_t^\epsilon(y))-f(\varphi_t(y))\right|dy\\&\quad+\|\xi_0\|_{L^\infty(\mathbb{T}^2)}\int_{\mathbb{T}^2\setminus C_\delta}\left|f(\varphi_t^\epsilon(y))\right|\mathrm{d}y\\&\quad+\|\xi_0\|_{L^\infty(\mathbb{T}^2)}\int_{\mathbb{T}^2\setminus C_\delta}|f(\varphi_t(y))|\,dy.\end{aligned}$$

Since $|f|\in L^1(\mathbb{T}^2)$ and φ_t^ϵ, φ_t are measure-preserving, absolute continuity of Lebesgue integral gives: for every $\delta'>0$ there exists $\delta>0$ such that, for every

$\epsilon > 0$, $t \in [0, T]$ and a.e. $\omega \in \Omega$:

$$\int_{\mathbb{T}^2 \setminus C_\delta} \left| f(\varphi_t^\epsilon(y)) \right| dy + \int_{\mathbb{T}^2 \setminus C_\delta} |f(\varphi_t(y))| \, dy < \delta'.$$

It remains to study the quantity

$$\int_{C_\delta} \left| f(\varphi_t^\epsilon(y)) - f(\varphi_t(y)) \right| dy \le \int_{\mathbb{T}^2} \left| f_\delta(\varphi_t^\epsilon(y)) - f_\delta(\varphi_t(y)) \right| \mathrm{d}y.$$

Since f_δ is continuous, one can argue as in (Brzeźniak et al. 2016, Proposition 6.2) to get

$$\mathbb{E}\left[\int_{\mathbb{T}^2} \left| f_\delta(\varphi_t^\epsilon(y)) - f_\delta(\varphi_t(y)) \right| dy \right] \to 0$$

as $\epsilon \to 0$, for every fixed $t \in [0, T]$ and in $L^p([0, T])$ for every finite p. Putting all together, the proof is complete. □

6 Back to 2D Euler Equations

In this section, we focus back to the *full* 2D Euler system (E)

$$\begin{cases} \mathrm{d}\xi_{\mathrm{L}}^\epsilon + u_{\mathrm{L}}^\epsilon \cdot \nabla \xi_{\mathrm{L}}^\epsilon \mathrm{d}t = -u_{\mathrm{S}}^\epsilon \cdot \nabla \xi_{\mathrm{L}}^\epsilon \mathrm{d}t, \\ d\xi_{\mathrm{S}}^\epsilon + u_{\mathrm{L}}^\epsilon \cdot \nabla \xi_{\mathrm{S}}^\epsilon \mathrm{d}t = -\epsilon^{-2} \xi_{\mathrm{S}}^\epsilon \mathrm{d}t + \epsilon^{-2} \mathrm{d}W_t, \\ \xi_{\mathrm{L}}^\epsilon|_{t=0} = \xi_0, \quad \xi_{\mathrm{S}}^\epsilon|_{t=0} = \xi_{\mathrm{S}}^{0,\epsilon}. \end{cases}$$

Recall that the Brownian motion $W(t, x)$ is given by

$$W(t, x) = \sum_{k \in \mathbb{N}} \theta_k(x) \beta_t^k,$$

where the coefficients θ_k satisfy assumption (A1):

$$\theta_k \in L_0^2(\mathbb{T}^2) \cap C^1(\mathbb{T}^2, \mathbb{R}),$$
$$\sum_{k \in \mathbb{N}} \|\nabla \theta_k\|_{L^\infty(\mathbb{T}^2, \mathbb{R}^2)} < \infty.$$

Let $\Theta^\epsilon(t, x) = \sum_{k \in \mathbb{N}} \theta_k(x) \eta_t^{\epsilon,k}$ be the Ornstein–Uhlenbeck process solution of

$$d\Theta^\epsilon = -\epsilon^{-2} \Theta^\epsilon \mathrm{d}t + \epsilon^{-2} \mathrm{d}W_t,$$

with initial condition $\xi_{\mathrm{S}}^{0,\epsilon}$. For simplicity, take $\xi_{\mathrm{S}}^{0,\epsilon}$ as in Sect. 3, so that Θ^ϵ is a stationary process, progressively measurable with respect to the filtration $\{\mathcal{F}_t\}_{t \ge 0}$, and

$\eta^{\epsilon,k}$ is independent of $\eta^{\epsilon,h}$ for $k \neq h$. Notice that the regularity of Θ^ϵ is the same of W. In particular, under assumption (A1), Θ^ϵ takes a.s. values in $C([0,T], W^{1,\infty}(\mathbb{T}^2))$ and

$$\mathbb{E}\left[\sup_{t\in[0,T]} \|\nabla\Theta^\epsilon\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^2)}\right] \leq C\epsilon^{-1}. \tag{16}$$

Define the difference process:

$$\zeta^\epsilon = \xi_{\mathrm{S}}^\epsilon - \Theta^\epsilon,$$

which solves the equation

$$\begin{gathered} \mathrm{d}\zeta^\epsilon + u_{\mathrm{L}}^\epsilon \cdot \nabla\zeta^\epsilon \mathrm{d}t = -\epsilon^{-2}\zeta^\epsilon \mathrm{d}t - u_{\mathrm{L}}^\epsilon \cdot \nabla\Theta^\epsilon \mathrm{d}t, \\ \zeta_0^\epsilon = 0. \end{gathered}$$

We consider also the following auxiliary process

$$\tilde{\zeta}_t^\epsilon = e^{\epsilon^{-2}t}\zeta_t^\epsilon,$$

which solves the same equation without damping:

$$\begin{gathered} d\tilde{\zeta}^\epsilon + u_{\mathrm{L}}^\epsilon \cdot \nabla\tilde{\zeta}^\epsilon \mathrm{d}t = -u_{\mathrm{L}}^\epsilon \cdot \nabla\Theta^\epsilon dt, \\ \tilde{\zeta}_0^\epsilon = 0. \end{gathered}$$

By (Kunita 1997, Theorem 6.1.6) the problem above admits formally an unique solution:

$$\tilde{\zeta}^\epsilon(t,x) = -\int_0^t (u_{\mathrm{L}}^\epsilon \cdot \nabla\Theta^\epsilon)(s, (\phi_{s,t}^\epsilon)^{-1}(x))\mathrm{d}s, \tag{17}$$

where ϕ^ϵ is defined by

$$d\phi_{s,t}^\epsilon(x) = u_{\mathrm{L}}^\epsilon(t, \phi_{s,t}^\epsilon(x))\mathrm{d}t, \quad \phi_{s,s}^\epsilon(x) = x.$$

Recalling the equality $\tilde{\zeta}_t^\epsilon = e^{\epsilon^{-2}t}\zeta_t^\epsilon$, the equation above becomes

$$\zeta^\epsilon(t,x) = -e^{-\epsilon^{-2}t}\int_0^t (u_{\mathrm{L}}^\epsilon \cdot \nabla\Theta^\epsilon)(s, \phi_{s,t}^{-1}(x))\mathrm{d}s.$$

By difference, we recover the small-scale vorticity $\xi_{\mathrm{S}}^\epsilon$ and we can plug it into the equation for the large-scale vorticity $\xi_{\mathrm{L}}^\epsilon$ to obtain:

$$\mathrm{d}\xi_{\mathrm{L}}^{\epsilon} + u_{\mathrm{L}}^{\epsilon} \cdot \nabla \xi_{\mathrm{L}}^{\epsilon} \mathrm{d}t = -(K * \Theta^{\epsilon}) \cdot \nabla \xi_{\mathrm{L}}^{\epsilon} dt + e^{-\epsilon^{-2} t} K * \left(\int_0^t (u_{\mathrm{L}}^{\epsilon} \cdot \nabla \Theta^{\epsilon})(s, \phi_{s,t}^{-1}(\cdot)) ds \right) \cdot \nabla \xi_{\mathrm{L}}^{\epsilon} \mathrm{d}t \qquad (18)$$

This is a (highly nonlinear) transport equation with random coefficients. It is worth noticing that we have obtained equation (18) above by a formal application of (17), somewhat in the same spirit of formal integration-by-parts performed when dealing with weak solutions of certain PDEs. Well-posedness, in the Lagrangian sense—that is, the analogous of Definition 3.3—of (18) is the content of the following:

Proposition 6.1 *For every $\epsilon > 0$, equation (18) admits a unique weakly progressively measurable solution ξ_L^{ϵ}, given by the transportation of the initial vorticity ξ_0 along the characteristics ψ^{ϵ} defined below. Moreover, the characteristics ψ^{ϵ} of the full Eq. (18) converge to the characteristics (12) as $\epsilon \to 0$ in the following sense:*

$$\mathbb{E}\left[\sup_{s \le T} \left\| \psi_s^{\epsilon} - \varphi_s \right\|_{L^1(\mathbb{T}^2, \mathbb{T}^2)} \right] \to 0.$$

Proof Consider the following system of characteristics:

$$\begin{cases} \mathrm{d}\psi_t^{\epsilon}(x) = u_{\mathrm{L}}^{\epsilon}(t, \psi_t^{\epsilon}(x))\mathrm{d}t + (K * \Theta^{\epsilon})(t, \psi_t^{\epsilon}(x))\mathrm{d}t \\ \qquad\qquad + (K * \zeta^{\epsilon})(t, \psi_t^{\epsilon}(x))\mathrm{d}t, \quad \psi_0^{\epsilon}(x) = x, \\ \xi_{\mathrm{L}}^{\epsilon}(t, x) = \xi_0((\psi^{\epsilon}{}_t)^{-1}(x)), \\ u_{\mathrm{L}}^{\epsilon}(t, x) = (K * \xi_{\mathrm{L}}^{\epsilon})(t, x), \\ \mathrm{d}\phi_{s,t}^{\epsilon}(x) = u_{\mathrm{L}}^{\epsilon}(t, \phi_{s,t}^{\epsilon}(x))dt, \quad \phi_{s,s}^{\epsilon}(x) = x, \\ \zeta^{\epsilon}(t, x) = -e^{-\epsilon^{-2} t} \int_0^t (u_{\mathrm{L}}^{\epsilon} \cdot \nabla \Theta^{\epsilon})(s, (\phi_{s,t}^{\epsilon})^{-1}(x))\mathrm{d}s. \end{cases} \qquad (C)$$

Notice that the only unknown of the system above is the characteristic ψ^{ϵ}, the other quantities being uniquely determined ω-wise by the former, see (Brzeźniak et al. 2016, Section 3). We prove path-by-path well-posedness of the system (C) in the class of flows of measure-preserving homeomorphisms. The argument is similar to that of the proof of (Brzeźniak et al. 2016, Theorem 3.4).

Let $\mathcal{M}_T$ be the space

$$\mathcal{M}_T = \Big\{ \psi : [0, T] \times \mathbb{T}^2 \to \mathbb{T}^2 \text{ measurable }, \; \psi_t \text{ is a measure-preserving homeomorphism for every } t \in [0, T] \Big\}.$$

The proof relies on a Picard iteration with *fixed* ω. Let $G : \mathcal{M}_T \mapsto \mathcal{M}_T$ be the map that at every ψ associates the solution $G(\psi)$ of the equation:

$$G(\psi)_t(x) = x + \int_0^t u(s, G(\psi)_s(x)) ds + \int_0^t (K * \Theta^{\epsilon})(s, G(\psi)_s(x))\mathrm{d}s$$

$$+ \int_0^t (K * \zeta)(s, G(\psi)_s(x))\mathrm{d}s,$$

where $u = u^\psi$, $\zeta = \zeta^\psi$ are computed from ψ and not from $G(\psi)$, so that the $G(\psi)$ is well defined for every $\psi \in \mathcal{M}_T$ (see Brzeźniak et al. 2016, Lemma 3.1 for the analogous result for Euler equations). For any $\psi, \psi' \in \mathcal{M}_T$, a computation similar to that of the proof of Proposition 4.1 yields the key inequality

$$\begin{aligned}\left\|G(\psi)_t - G(\psi')_t\right\|_{L^1(\mathbb{T}^2,\mathbb{T}^2)} \lesssim & \left(1 + \sup_{s\le T}\left\|\nabla\Theta_s^\epsilon\right\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^2)}\right) \\ & \times \int_0^t \gamma\left(\left\|G(\psi)_s - G(\psi')_s\right\|_{L^1(\mathbb{T}^2,\mathbb{T}^2)}\right) ds \\ & + \left(1 + \sup_{s\le T}\left\|\nabla\Theta_s^\epsilon\right\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^2)}\right) \\ & \times \int_0^t \gamma\left(\left\|\psi_s - \psi'_s\right\|_{L^1(\mathbb{T}^2,\mathbb{T}^2)}\right)\mathrm{d}s,\end{aligned}$$

which guarantees the a.s. convergence of the Picard iteration towards a solution of the system (C) on the time interval $[0, T_1]$, where $0 < T_1 \le T$ may depend on ω. However, for any fixed ω, one can iterate this procedure with the same time step T_1, to obtain existence on $[0, T]$ after $N = N(\omega)$ iterations of the argument. In addition, having care to initialize the iteration scheme with a $\mathcal{F}_0$ measurable random element of $\mathcal{M}_T$ (take for instance $\psi_t^0(x) = x$ for every t), we also obtain progressively measurability of the solution so constructed with respect to the filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$. Uniqueness is obtained applying the same Picard scheme to two solutions $\psi, \psi' \in \mathcal{M}_T$. We omit the remaining details, which are contained in Brzeźniak et al. (2016).

Let us now investigate the convergence of characteristics $\psi^\epsilon \to \varphi$. The proof is the same as Proposition 4.1. We do not repeat it here, and we limit ourselves to notice that since $\sup_{t\in[0,T]} \|\xi_L^\epsilon(t)\|_{L^\infty(\mathbb{T}^2)} \le \|\xi_0\|_{L^\infty(\mathbb{T}^2)}$, we have

$$\sup_{s\le T}\left\|\zeta_s^\epsilon\right\|_{L^1(\mathbb{T}^2)} \lesssim \|\xi_0\|_{L^\infty(\mathbb{T}^2)}\epsilon^2 \sup_{s\le T}\left\|\nabla\Theta_s^\epsilon\right\|_{L^\infty(\mathbb{T}^2,\mathbb{R}^2)}.$$

By (16), the expected value of this quantity is infinitesimal as $\epsilon \to 0$ and therefore does not affect the argument of Proposition 4.1. □

In virtue of the previous proposition, we deduce the analogous of Theorem 5.1 and Corollary 4.9 for the *full* 2D Euler system (E), that is Theorem 1.1 in Introduction, whose precise formulation is the following:

Theorem 6.2 *Assume (A1), and let ξ_L^ϵ be the large-scale process solution of* (E) *in the sense of Proposition 6.1, ξ_L be the solution of* (10)*. Then, ξ_L^ϵ converges as $\epsilon \to 0$ to ξ_L in the following sense: for every $f \in L^1(\mathbb{T}^2)$:*

$$\mathbb{E}\left[\left|\int_{\mathbb{T}^2} \xi_L^\epsilon(t,x) f(x)\,\mathrm{d}x - \int_{\mathbb{T}^2} \xi_L(t,x) f(x)\,\mathrm{d}x\right|\right] \to 0$$

as $\epsilon \to 0$, *for every fixed* $t \in [0, T]$ *and in* $L^p([0, T])$ *for every finite* p. *Moreover, the large-scale velocity process* $u_L^\epsilon = K * \xi_L^\epsilon$ *converges towards* $u_L = K * \xi_L$ *as* $\epsilon \to 0$, *in mean value, as variables in* $C([0, T], L^1(\mathbb{T}^2, \mathbb{R}^2))$.

Funding Open access funding provided by Scuola Normale Superiore within the CRUI-CARE Agreement. (IST Austria).

Appendix A. Convergence of Weak Solutions

Throughout the paper, we have used the Lagrangian formulation of Euler equations and, more generally, of transport-type equations. This point of view turns out to be really effective to investigate the convergence of the large-scale component of the system (E) as $\epsilon \to 0$, since the nonlinear term in the equation of characteristics has been widely studied before.

In this section, we aim to link the Lagrangian point of view on Euler equations and other equations of transport type with the analytically weak point of view.

To fix the ideas, take first 2D Euler equations in vorticity form (1). Both Lagrangian formulation and weak formulation aim to give a notion of solution to (1) which does not require the regularity needed for classical solution. In the case of Lagrangian formulation, the space derivative of the solution is formally cancelled out by the composition with the characteristics:

$$\partial_t \xi_t(\varphi_t(x)) = 0.$$

In the weak formulation, the derivatives of the solution are formally eliminated by an integration-by-parts formula for the product of the solution against regular test functions: to be precise, a weak solution of (1) is given by a function $\xi \in L^\infty([0, T], L^\infty(\mathbb{T}^2))$ such that, for every test function $f \in C^1(\mathbb{T}^2)$, it holds for every $t \in [0, T]$:

$$\int_{\mathbb{T}^2} \xi_t(x) f(x) \mathrm{d}x = \int_{\mathbb{T}^2} \xi_0(x) f(x) \mathrm{d}x + \int_0^t \left(\int_{\mathbb{T}^2} \xi_s(x) (K * \xi_s)(x) \cdot \nabla f(x) \mathrm{d}x \right) \mathrm{d}s.$$

In the case of 2D Euler equations in vorticity form, the Lagrangian point of view and the analytically weak point of view are equivalent, that is, every Lagrangian solution of (1) is also a weak solution, and every weak solution of (1) is given by the transportation of the initial datum ξ_0 along characteristics. The proof of this classical fact under the assumption $\xi_0 \in L_0^\infty(\mathbb{T}^2)$ can be found in Marchioro and Pulvirenti (1994).

In Brzeźniak et al. (2016) a similar statement is proved for the limiting process (10):

$$\mathrm{d}\xi_t + u_t \cdot \nabla \xi_t dt = -\sum_{k\in\mathbb{N}} \sigma_k \cdot \nabla \xi_t \circ \mathrm{d}\beta_t^k,$$

with $u_t = K * \xi_t$. In this case, for every initial datum $\xi_0 \in L_0^\infty(\mathbb{T}^2)$, the Lagrangian formulation of (10), which has been used in the present work, is equivalent to the following distributional formulation (see Brzeźniak et al. 2016, Theorem 2.14 and Proposition 5.3).

Definition A.1 A weakly progressively measurable process $\xi \in L^\infty(\Omega \times [0, T] \times \mathbb{T}^2)$ is said to be a L^∞ *distributional solution* to (10) if for every test function $f \in C^\infty(\mathbb{T}^2)$ it holds $\mathbb{P}$-a.s.: for every $t \in [0, T]$

$$\begin{aligned}
\int_{\mathbb{T}^2} \xi_t(x) f(x)\mathrm{d}x &= \int_{\mathbb{T}^2} \xi_0(x) f(x)\mathrm{d}x + \int_0^t \left(\int_{\mathbb{T}^2} \xi_s(x)(K * \xi_s)(x) \cdot \nabla f(x)\mathrm{d}x \right) \mathrm{d}s \\
&\quad + \sum_{k\in\mathbb{N}} \int_0^t \left(\int_{\mathbb{T}^2} \xi_s(x)\sigma_k(x) \cdot \nabla f(x) dx \right) \mathrm{d}\beta_s^k \\
&\quad - \frac{1}{2} \sum_{k\in\mathbb{N}} \int_0^t \left(\int_{\mathbb{T}^2} \xi_s(x)[(\sigma_k(x) \cdot \nabla)\sigma_k(x)] \cdot \nabla f(x)\mathrm{d}x \right) \mathrm{d}s \\
&\quad - \frac{1}{2} \sum_{k\in\mathbb{N}} \int_0^t \left(\int_{\mathbb{T}^2} \xi_s(x)\, div[(\sigma_k \cdot \nabla)\sigma_k](x) f(x)\mathrm{d}x \right) \mathrm{d}s \\
&\quad + \frac{1}{2} \sum_{k\in\mathbb{N}} \int_0^t \left(\int_{\mathbb{T}^2} \xi_s(x)\, tr[\sigma_k \sigma_k^* \nabla^2 f](x)\mathrm{d}x \right) \mathrm{d}s.
\end{aligned}$$

Following Brzeźniak et al. (2016), we define an analogous notion of L^∞ distributional solution to (18). Recall the definition of the Ornstein–Uhlenbeck process:

$$\Theta^\epsilon(t, x) = \sum_{k\in\mathbb{N}} \theta_k(x)\eta_t^{\epsilon,k}.$$

Definition A.2 A weakly progressively measurable process $\xi^\epsilon \in L^\infty(\Omega \times [0, T] \times \mathbb{T}^2)$ is said to be a L^∞ *distributional solution* to (18) if for every test function $f \in C^\infty(\mathbb{T}^2)$ it holds $\mathbb{P}$-a.s.: for every $t \in [0, T]$

$$\begin{aligned}
\int_{\mathbb{T}^2} \xi_t^\epsilon(x) f(x)\mathrm{d}x &= \int_{\mathbb{T}^2} \xi_0(x) f(x)\mathrm{d}x + \int_0^t \left(\int_{\mathbb{T}^2} \xi_s^\epsilon(x) u_s^\epsilon(x) \cdot \nabla f(x)\mathrm{d}x \right) \mathrm{d}s \\
&\quad + \int_0^t \left(\int_{\mathbb{T}^2} \xi_s^\epsilon(x)(K * \Theta_s^\epsilon)(x) \cdot \nabla f(x)\mathrm{d}x \right) \mathrm{d}s \\
&\quad + \int_0^t \left(\int_{\mathbb{T}^2} \xi_s^\epsilon(x)(K * \zeta_s^\epsilon)(x) \cdot \nabla f(x)\mathrm{d}x \right) \mathrm{d}s,
\end{aligned}$$

where $u^\epsilon(t,x) = (K * \xi^\epsilon)(t,x)$ and ζ^ϵ is given by

$$\zeta^\epsilon(t,x) = -e^{-\epsilon^{-2}t}\int_0^t (u^\epsilon\cdot\nabla\Theta^\epsilon)(s,\phi_{s,t}^{-1}(x))\mathrm{d}s,$$
$$\mathrm{d}\phi_{s,t}^\epsilon(x) = u^\epsilon(t,\phi_{s,t}^\epsilon(x))\mathrm{d}t, \quad \phi_{s,s}^\epsilon(x) = x.$$

As for equations (1) and (10), the notion of L^∞ distributional solution to (18) is indeed equivalent to the notion of Lagrangian solution used throughout the paper. This is the content of the forthcoming:

Proposition A.3 *The unique Lagrangian solution ξ^ϵ to (18) given by Proposition 6.1 is also a L^∞ distributional solution. Conversely, every L^∞ distributional solution to (18) is also a Lagrangian solution.*

Proof *Step 1*. We first prove that the Lagrangian solution to (18) is also a L^∞ distributional solution. Let ψ^ϵ be the unique stochastic flow of homeomorphism solution of the system of characteristics (C). By the representation formula $\xi_t^\epsilon = \xi_0\circ(\psi_t^\epsilon)^{-1}$, one immediately has $\xi^\epsilon \in L^\infty(\Omega\times[0,T]\times\mathbb{T}^2)$. By the a.s. measure-preserving property of ψ^ϵ, one gets, for any $f \in L^1(\mathbb{T}^2)$:

$$\int_{\mathbb{T}^2}\xi_t^\epsilon(x)f(x)\mathrm{d}x = \int_{\mathbb{T}^2}\xi_0(x)f(\psi_t^\epsilon(x))\mathrm{d}x. \tag{19}$$

Since for every $x\in\mathbb{T}^2$ the process $t\mapsto\psi_t^\epsilon(x)$ is progressively measurable, one can deduce from (19) that ξ^ϵ is weakly progressively measurable.

Let now $f\in C^\infty(\mathbb{T}^2)$ be a given test function. Define

$$v_s^\epsilon(x) = u_s^\epsilon(x) + (K*\Theta_s^\epsilon)(x) + (K*\zeta_s^\epsilon)(x).$$

Using

$$df(\psi_t^\epsilon(x)) = v_t^\epsilon(\psi_t^\epsilon(x))\cdot\nabla f(\psi_t^\epsilon(x))\mathrm{d}t,$$

multiplying per ξ_0, integrating in time and space, and using (19) one obtains that ξ^ϵ is a L^∞ distributional solution to (18).

Step 2. We prove that every L^∞ distributional solution to (18) is also a Lagrangian solution, *i.e.* it is given by the transportation of the initial vorticity ξ_0 along characteristics.

A weakly progressively measurable process $\xi^\epsilon\in L^\infty([0,T]\times\mathbb{T}^2\times\Omega)$ being a L^∞ distributional solution to (18) corresponds to: for every test function $f\in C^\infty(\mathbb{T}^2)$ it holds $\mathbb{P}$-a.s.: for every $t\in[0,T]$

$$\int_{\mathbb{T}^2}\xi_t^\epsilon(x)f(x)\mathrm{d}x = \int_{\mathbb{T}^2}\xi_0(x)f(x)\mathrm{d}x + \int_0^t\left(\int_{\mathbb{T}^2}\xi_s^\epsilon(x)v_s^\epsilon(x)\cdot\nabla f(x)\mathrm{d}x\right)\mathrm{d}s. \tag{20}$$

Let $\rho \in C^\infty(\mathbb{R}^2)$ be a non-negative even function, supported in $[-1,1]^2$, with $\int_{\mathbb{R}^2} \rho(x)\mathrm{d}x = 1$. For every $\delta \in (0, 1/2)$, denote $\rho_\delta(x) = \delta^{-2}\rho(x/\delta)$ and

$$\xi_t^{\epsilon,\delta}(x) = \int_{\mathbb{R}^2} \rho_\delta(x-y)\xi_t^\epsilon(y)dy.$$

In the integral above, the function $\xi_t^\epsilon : \mathbb{T}^2 \to \mathbb{R}$ is interpreted as a periodic function $\xi_t^\epsilon : \mathbb{R}^2 \to \mathbb{R}$ on the full space. The mollified vorticity $\xi_t^{\epsilon,\delta}$ is smooth and periodic in space; therefore, the process $\xi^{\epsilon,\delta}$ has a.s. trajectories in

$$\xi^{\epsilon,\delta} \in L^\infty([0,T], C^\infty(\mathbb{T}^2)).$$

Using that ξ^ϵ is a L^∞ distributional solution to (18) in the equivalent formulation (20) with $f = \rho_\delta(x - \cdot)$, one has for every fixed $x \in \mathbb{T}^2$ the a.s. property: for every $t \in [0,T]$

$$\xi_t^{\epsilon,\delta}(x) = (\xi_0 * \rho_\delta)(x) + \int_0^t \left(\int_{\mathbb{T}^2} (\xi_s^\epsilon(y) v_s^\epsilon(y) \cdot \nabla \rho_\delta(x-y) dy \right) \mathrm{d}s.$$

Notice that, by a.s. space regularity of the process $\xi^{\epsilon,\delta}$, one can find a full-measure set $\Omega' \subset \Omega$ such that for every $\omega \in \Omega'$ the property above holds simultaneously for every $x \in \mathbb{T}^2$. Arguing as in Brzeźniak et al. (2016), for ψ^ϵ solving

$$d\psi_t^\epsilon(x) = v^\epsilon(t, \psi_t^\epsilon(x))\mathrm{d}t, \quad \psi_0^\epsilon(x) = x,$$

one can prove a.s. the following: for every $x \in \mathbb{T}^2$ and $t \in [0,T]$

$$\xi_t^{\epsilon,\delta}(\psi_t^\epsilon(x)) = (\xi_0 * \rho_\delta)(x) + \int_0^t \left[v_s^\epsilon \cdot \nabla, *\rho_\delta \right] \xi_s^\epsilon(\psi_s^\epsilon(x))\mathrm{d}s. \tag{21}$$

The commutator above is defined for fixed divergence-free $v \in W^{1,p}(\mathbb{T}^2, \mathbb{R}^2)$, $p \in [1, \infty)$, and $w \in L^\infty(\mathbb{T}^2)$ by

$$[v \cdot \nabla, *\rho_\delta]w = v \cdot \nabla(\rho_\delta * w) - \rho_\delta * (v \cdot \nabla w),$$

and satisfies (see Brzeźniak et al. 2016, Lemma 5.2)

$$\lim_{\delta \to 0} [v \cdot \nabla, *\rho_\delta]w = 0 \text{ in } L^p(\mathbb{T}^2), \tag{22}$$

$$\|[v \cdot \nabla, *\rho_\delta]w\|_{L^p(\mathbb{T}^2)} \lesssim \|\nabla v\|_{L^p(\mathbb{T}^2)} \|w\|_{L^\infty(\mathbb{T}^2)}. \tag{23}$$

By measure-preserving property of ψ^ϵ, integrating (21) in space yields

$$\int_{\mathbb{T}^2} |\xi_t^{\epsilon,\delta}(\psi_t^\epsilon(x)) - (\xi_0 * \rho_\delta)(x)|\mathrm{d}x \lesssim \int_0^t \int_{\mathbb{T}^2} | \left[v_s^\epsilon \cdot \nabla, *\rho_\delta \right] \xi_s^\epsilon(x)|\mathrm{d}x\mathrm{d}s.$$

Using (22) and (23) with $p = 1$, the bound

$$\|\nabla v_s^\epsilon\|_{L^1(\mathbb{T}^2)} \lesssim \|\xi_s^\epsilon\|_{L^\infty(\mathbb{T}^2)} + \|\Theta_s^\epsilon\|_{L^\infty(\mathbb{T}^2)} + \|\zeta_s^\epsilon\|_{L^\infty(\mathbb{T}^2)},$$

and Lebesgue dominated convergence theorem, one obtains a.s. the convergence: for every fixed $t \in [0, T]$, as $\delta \to 0$:

$$\xi_t^{\epsilon,\delta}(\psi_t^\epsilon(\cdot)) - (\xi_0 * \rho_\delta) \to 0 \text{ in } L^1(\mathbb{T}^2).$$

By the fact of ξ^ϵ being a.s. with trajectories in $L^\infty(\mathbb{T}^2)$ and ψ^ϵ being a.s. measure-preserving, one also has a.s. the convergence: for every fixed $t \in [0, T]$, as $\delta \to 0$:

$$\xi_t^{\epsilon,\delta}(\psi_t^\epsilon(\cdot)) - \xi_t^\epsilon(\psi_t^\epsilon(\cdot)) \to 0 \text{ in } L^1(\mathbb{T}^2),$$

and similarly

$$(\xi_0 * \rho_\delta) - \xi_0 \to 0 \text{ in } L^1(\mathbb{T}^2).$$

We remark that the previous statement asserts the possibility of finding a full-measure set $\Omega'' \subset \Omega$ such that for every $\omega \in \Omega''$ the convergences above hold for every fixed $t \in [0, T]$. Putting all together, we finally obtain a.s. the identity: for every $t \in [0, T]$, $\xi_t^\epsilon(\psi_t^\epsilon(\cdot)) = \xi_0$ as variables in $L^1(\mathbb{T}^2)$, that is for Lebesgue-a.e. $x \in \mathbb{T}^2$. By boundedness, the identity can be understood as variables in $L^\infty(\mathbb{T}^2)$ as well. □

Proposition A.3 gives well-posedness of (18) in distributional formulation: indeed, the Lagrangian solution given by Proposition 6.1 is a L^∞ distributional solution, giving existence; for uniqueness, it suffices to invoke uniqueness of characteristics and the fact that every L^∞ distributional solution is also a Lagrangian solution. In terms of convergence, one can therefore restate Theorem 6.2 with the L^∞ distributional notion of solution.

References

Assing, S., Flandoli, F., Pappalettera, U.: Stochastic model reduction: convergence and applications to climate equations, arXiv:2010.05216, (2020)

Barbato, D., Bessaih, H., Ferrario, B.: On a stochastic Leray-α model of Euler equations. Stochastic Process. Appl. **124**(1), 199–219 (2014)

Beck, L., Flandoli, F., Gubinelli, M., Maurelli, M.: Stochastic ODEs and stochastic linear PDEs with critical drift: regularity, duality and uniqueness. *Electr. J. of Probab.*, page arXiv:1401.1530, (2014)

Bianchi, L.A., Flandoli, F.: Stochastic Navier–Stokes equations and related models. Milan J. Math. **88**(1), 225–246 (2020)

Bianchi, L.A.: Uniqueness for an inviscid stochastic dyadic model on a tree. Electron. Commun. Probab. **18**(8), 12 (2013)

Boffetta, G., Ecke, R.E.: Two-dimensional turbulence. Ann. Rev. Fluid Mech. **44**(1), 427–451 (2012)

Brzeźniak, Z., Capiński, M., Flandoli, F.: A convergence result for stochastic partial differential equations. Stochastics **24**(4), 423–445 (1988)

Brzeźniak, Z., Capiński, M., Flandoli, F.: Stochastic partial differential equations and turbulence. Math. Models Methods Appl. Sci. **1**(1), 41–59 (1991)

Brzeźniak, Z., Capiński, M., Flandoli, F.: Stochastic Navier-Stokes equations with multiplicative noise. Stochastic Anal. Appl. **10**(5), 523–532 (1992)
Brzeźniak, Z., Flandoli, F., Maurelli, M.: Existence and uniqueness for stochastic 2D Euler flows with bounded vorticity. Arch. Ration. Mech. Anal. **221**(1), 107–142 (2016)
Chae, D., Constantin, P., Wu, J.: Inviscid models generalizing the two-dimensional Euler and the Surface Quasi-geostrophic equations. Arch. Ration. Mech. Anal. **202**, 35–62 (2011)
Cheskidov, A., Holm, D.D., Olson, E., Titi, E.S.: On a Leray-α model of turbulence. Proc. R. Soc. A. **461**, 629–649 (2005)
Cotter, C.J., Gottwald, G.A., Holm, D.D.: Stochastic partial differential fluid equations as a diffusive limit of deterministic lagrangian multi-time dynamics. Proc. R. Soc. A. **473**, 20170388 (2017)
Crisan, D., Flandoli, F., Holm, D.D.: Solution properties of a 3d stochastic Euler fluid equation. J. Nonlinear Sci. **29**(3), 813–870 (2019)
Flandoli, F., Gubinelli, M., Priola, E.: Well-posedness of the transport equation by stochastic perturbation. Invent. Math. **180**(1), 1–53 (2010)
Flandoli, F.: *Random perturbation of PDEs and fluid dynamic models*. Lecture Notes in Mathematics. Springer, Heidelberg, 2011. Lectures from the 40th Probability Summer School held in Saint-Flour, École d'Été de Probabilités de Saint-Flour (2010)
Flandoli, F., Maurelli, M., Neklyudov, M.: Noise prevents infinite stretching of the passive field in a stochastic vector advection equation. J. Math. Fluid Mech. **16**(4), 805–822 (2014)
Franzke, C.L.E., Oliver, M., Rademacher, J.D.M., Badin, G.: Multi-scale methods for geophysical flows. In: Eden, C., Iske, A. (eds.) Energy Transfer in Atmosphere and Ocean, pp. 1–52. Springer, Berlin (2019)
Franzke, C., Majda, A.J.: Low-order stochastic mode reduction for a prototype atmospheric GCM. J. Atmos. Sci. **63**(2), 457–479 (2006)
Franzke, C., Majda, A.J., Vanden-Eijnden, E.: Low-order stochastic mode reduction for a realistic barotropic model climate. J. Atmos. Sci. **62**(6), 1722–1745 (2005)
Gay-Balmaz, F., Holm, D.D.: Stochastic geometric models with non-stationary spatial correlations in lagrangian fluid flows. J. Nonlinear Sci. **28**(3), 873–904 (2018)
Hofmanová, M., Leahy, J., Nilssen, T.: On the Navier-Stokes equation perturbed by rough transport noise. J. Evol. Equ. **19**, 203–247 (2019)
Hofmanova, M., Leahy, James-Michael, N., Torstein: On a rough perturbation of the navier-stokes system and its vorticity formulation, (2019)
Holm, D.D.: Variational principles for stochastic fluid dynamics. Proc. R. Soc. A. **471**, 20140963 (2015)
Ikeda, N., Watanabe, S.: *Stochastic Differential Equations and Diffusion Processes*. North-Holland Mathematical Library. North-Holland Publishing Co., second edition, (1989)
Jain, A., Timofeyev, I., Vanden-Eijnden, E.: Stochastic mode-reduction in models with conservative fast sub-systems. Commun. Math. Sci. **13**(2), 297–314 (2015)
Jia, C., Zhao, G.: Moderate maximal inequalities for the Ornstein-Uhlenbeck process. Proc. Am. Math. Soc. (2020). https://doi.org/10.1090/proc/14804
Kraichnan, R.H.: Inertial ranges in two-dimensional turbulence. Phys. Fluids **10**(7), 1417–1423 (1967)
Kunita, H.: *Stochastic Flows and Stochastic Differential Equations*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge (1997)
Majda, A.J., Timofeyev, I., Eijnden, E.V.: A mathematical framework for stochastic climate models. Commun. Pure Appl. Math. **54**(8), 891–974 (2001)
Marchioro, C., Pulvirenti, M.: Mathematical theory of incompressible nonviscous fluids. Applied Mathematical Sciences, vol. 96. Springer, New York (1994)
Maurelli, M.: Wiener chaos and uniqueness for stochastic transport equation. C. R. Math. Acad. Sci. Paris **349**(11–12), 669–672 (2011)
Mikulevicius, R., Rozovskii, B.L.: Stochastic Navier–Stokes equations for turbulent flows. SIAM J. Math. Anal. **35**(5), 1250–1310 (2004)
Mikulevicius, R., Rozovskii, B.L.: Global l 2 -solutions of stochastic Navier–Stokes equations. Ann. Probab. **33**(1), 137–176 (2005)
Penland, C., Matrosova, L.: A balance condition for stochastic numerical models with application to the El Niño-Southern oscillation. J. Climate **7**(9), 1352–1372 (1994)
Pisier, G.: *Martingales in Banach Spaces*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge (2016)

Tessitore, G., Zabczyk, J.: Wong–Zakai approximations of stochastic evolution equations. J. Evolut. Equa. **6**(4), 621–655 (2006)

Twardowska, K.: Approximation theorems of Wong-Zakai type for stochastic differential equations in infinite dimensions. *Dissertationes Math. (Rozprawy Mat.)*, 325, (1993)

Rudin, W.: *Real and complex analysis*. MGH, (1970)

Wong, E., Zakai, M.: On the convergence of ordinary integrals to stochastic integrals. Ann. Math. Statist. **36**(5), 1560–1564 (1965)