Comparison of Stochastic Parametrization Schemes Using Data Assimilation on Triad Models

Alexander Lobbe, Dan Crisan, Darryl Holm, Etienne Mémin, Oana Lang, and Bertrand Chapron

Abstract In recent years, stochastic parametrizations have been ubiquitous in modelling uncertainty in fluid dynamics models. One source of model uncertainty comes from the coarse graining of the fine-scale data and is in common usage in computational simulations at coarser scales. In this paper, we look at two such stochastic parametrizations: the Stochastic Advection by Lie Transport (SALT) parametrization introduced by Holm (Proc A 471(2176):20140963, 19, 2015) and the Location Uncertainty (LU) parametrization introduced by Mémin (Geophys Astrophys Fluid Dyn 108(2):119–146, 2014). Whilst both parametrizations are available for full-scale models, we study their reduced order versions obtained by projecting them on a complex vector Fourier mode triad of eigenfunctions of the curl. Remarkably, these two parametrizations lead to the same reduced order model, which we term the *helicity-preserving stochastic triad* (HST). This reduced order model is then compared with an alternative model which preserves the energy of the system, and which is termed the *energy preserving stochastic triad* (EST). These low-dimensional models are ideal benchmark models for testing new Data Assimilation algorithms: they are easy to implement, exhibit diverse behaviours depending on the choice of the coefficients and come with natural physical properties such as the conservation of energy and helicity.

A. Lobbe (✉) · D. Crisan · D. Holm · O. Lang
Imperial College London, Mathematics, London, UK
e-mail: alex.lobbe@imperial.ac.uk

E. Mémin
Campus Universitaire de Beaulieu, Inria - Institut National de Recherche en Sciences et Technologies du Numérique, Rennes, France

B. Chapron
Ifremer – Institut Français de Recherche pour l'Exploitation de la Mer, Plouzané, France

1 Introduction

The introduction of stochasticity in fluid dynamics has recently been the subject of intense research effort. This approach involves using random processes to model, for example, unresolved scales, or to take into account neglected physical effects. A stochastic formulation for the fluid flow introduces a probabilistic basis for modelling unresolved scales. This is different from the deterministic causal modelling which is difficult to achieve in practice due, for instance, to unknown initial conditions. In addition, some phenomena such as energy backscattering are directly accessible as stochastic processes. Another usage of stochastic modelling is to generate ensembles of realizations of the model. This facilitates the analysis of model uncertainty quantification for different low-resolution computational simulations and their usage to approximate the true state of the fluid, instead of using a single high-resolution numerical simulation.

Some stochastic schemes have been proposed in the literature by considering a variety of ad-hoc perturbations. However, a principled approach is desirable. The formulation of stochastic dynamical systems based on physical principles has recently been proposed in various settings. For a review and classification of approaches to stochastic parameterisation based on physical principles, see [4]. The present work treats two additional new approaches. The first one, called stochastic advection by Lie transport (SALT) relies on the variational principle for fluid dynamics [16]. The second one, called modelling under location uncertainty (LU) is derived from Newton's principle [23]. Both frameworks introduce stochasticity into the Lagrangian specification of the flow field, rather than directly into the Eulerian frame.

In the deterministic case, it is known that three-dimensional fluid flows may trigger a cascade of dynamics across multiple length and time scales. This multiscale behavior poses considerable challenges in the computational simulation using standard Navier-Stokes equations (see e.g. [7], [24], [5]). When modelling turbulence numerically, specialised discretisation methods are needed to decompose the underlying partial differential equations into a very large number of ordinary differential equations. Alternative approaches have been introduced where the Navier-Stokes dynamics in the Fourier space is mimicked using a finite number of variables, say $u^1, u^2, \ldots, u^N$. The Fourier space is divided into N shells, and each shell s_i comprises the set of wave vectors $\underline{s}$ with magnitude $|\underline{s}| \in (s_0 2^i, s_0 2^{i+1})$. Each u^i satisfies an ODE and it represents the magnitude of the velocity field on a length scale of s_i^{-1} ([15], [7]). The quadratic nonlinearity in the Navier-Stokes equations produces triads of interacting vector Fourier modes within each shell. Shell models involving multiple triads have had considerable success in modelling energy and helicity cascades, as well as modelling intermittency in chaotic dynamical systems [6, 10, 9]. Simplified shell models with only a few triads date back to the 1970s and have provided major insight into fluid modal interaction. Even the dynamical system representation of Euler's fluid equations on a single triad has been quite insightful, see e.g. [28, 29].

More recently, the problem of correctly parameterizing effects of small-scale physical processes together with the need for probabilistic ensemble forecasting and uncertainty quantification has led to modern stochastic approaches in the study of turbulence using reduced order *shell models*. In this work we will explore reduced order models for SALT and LU models obtained by projecting onto helical basis functions [6, 28, 29]. These helical basis functions, defined as eigenfunctions of the curl operator, enable one to construct reduced order stochastic models of fluid flow with a simplified nonlinear interaction. As we will see, under projection onto the basis of helical triad modes, both LU and SALT result in the *same* reduced order model and this projected model conserves helicity, but it does not conserve energy. Because of this coincidence in projecting the SALT and LU models onto the helical basis, a second reduced order scheme with a strong energy conservation property inspired by [17] and known as the *energy preserving stochastic triad* (EST) model will be proposed for comparison.

While the EST model is not of transport type, it will provide comparison between two different classes of stochastic dynamical systems. The two classes treated here are (1) the *helicity preserving stochastic triad* (HST) (comprising both LU and SALT on the helical basis) and (2) the *energy preserving stochastic triad* (EST) of [17] projected onto the helical basis. The solution behaviour of the HST model will be compared to that for the EST model for several data assimilation objectives formulated on the helical triad modes. For classical deterministic models one obtains a system of ordinary differential equations. However, for stochastic dynamics a set of stochastic differential equations (SDEs) is obtained [8, 14, 25].

The goal of the data assimilation procedure in this context is twofold: firstly, it is used to calibrate the uncertainty of the model (the amplitude of the noise). Secondly, once the calibration is complete, the particle filtering methodology can be used to reduce the uncertainty. We want the distribution of the fluctuations to be properly approximated. In the absence of stochasticity, all particles would go in the same direction and the initial spread would rapidly disappear because of the hyperbolic character of the model. In the absence of a reasonable spread, the particle filter methodology will eventually collapse. For this reason, we need to introduce stochasticity into the system that correctly characterises the fluctuation dynamics. In particular, we want to find the type of noise amplitude and the stochastic parameters for which the distribution of the output samples is reasonably *uniform*.

Structure of the Paper In Sect. 2 we introduce the triad models for incompressible flows modelled by the Euler equation in its deterministic and stochastic form. To this end, we introduce the stochastic parametrisation paradigms. Building upon these models for the 3D Euler equation, we then present reduced order triad models derived from the original equations. The derivation follows the classical approach for triad models from the literature, that has been successfully employed in the deterministic case. Our full derivations, complete also for the stochastically parametrised models, can be found in Appendix 2. The Data Assimilation experiments are carried out in Sect. 3. We first briefly explain the standard particle filter methodology, and then in Sect. 3.1 we present the findings of our numerical studies. In particular, Sect. 3.1 presents the results of the main numerical studies in this work. These are:

- The model realisations of the stochastic models which we propose in this work are presented in Sect. 3.1.3 for different realisations of the noise. Further, we numerically confirm here the physical conservation properties of the models that motivated their theoretical design.
- The model statistics for a large number of stochastic realisations are presented in Sect. 3.1.4. Crucially, the evolution of the model statistics in time reveals differences between the two stochastic models on the level of the individual triad energies which go beyond the conservation properties. Moreover, we show here the (in-)stability of either model with respect to large noise parameters.
- In Sect. 3.1.5 we exhibit the results of the data assimilation performed using the transition kernels derived from the theoretical models developed in this paper. We show here that using the stochastic transition kernels associated with our proposed stochastic models improves the particle filtering procedure and produces efficient ensemble evolutions that are well-suited for data assimilation purposes.

In Sect. 4 we describe our conclusions on this topic. We conclude the paper with a number of appendices: in Appendix 1 one can find a list of notations and standard identities, in Appendix 2 we present a detailed derivation of shell models (deterministic and stochastic), in Appendix 3 we introduce some supplementary numerics related to the noise amplitude calibration.

Code Availability The code corresponding to the numerical experiments in this paper is archived in [22]. The GitHub repository is located at https://github.com/alobbe/stochastic-triads.

2 Reduced Order Models for Incompressible Fluids

2.1 Reduced Order Models for the 3D Euler Equation

The 3D Euler equations model incompressible inviscid fluid dynamics. These equations may be written by using the Leray operator $\mathcal{P}$ to project onto the divergence-free part of its operand as

$$
\begin{aligned}
\frac{\partial \mathbf{v}}{\partial t} &= \mathcal{P}\Big(\mathbf{v} \times \mathrm{curl}\mathbf{v}\Big) \\
&= \mathcal{P}\left(\frac{\delta E}{\delta \mathbf{v}} \times \frac{\delta C}{\delta \mathbf{v}}\right)
\end{aligned}
$$

with conserved Energy $E(\mathbf{v}) = \displaystyle\int_{\mathbb{R}^3} \frac{1}{2}\mathcal{P}\mathbf{v} \cdot \mathbf{v}\, d^3x$

and conserved helicity $C(\mathbf{v}) = \dfrac{1}{2}\displaystyle\int_{\mathbb{R}^3} \mathbf{v} \cdot \mathrm{curl}\,\mathbf{v}\, d^3x$

$$(1)$$

where $\delta/\delta\mathbf{v}$ represents variational derivative with respect to the fluid velocity $\mathbf{v}$.

Following [9, 10] we use a Galerkin expansion in orthogonal *vector* modes that are eigenfunctions of the curl operator. Assume the fluid is contained in a periodic box $\mathcal{D} \subset \mathbb{R}^3$ of side length $L > 0$. Then the velocity $\mathbf{v}$ and vorticity $\boldsymbol{\omega} := \nabla \times \mathbf{v}$ may be expanded in circularly polarised or helical modes $\mathbf{h}_\pm(\mathbf{k}) \exp(i\mathbf{k} \cdot \mathbf{x})$, with the wave vectors $\mathbf{k} \in \mathcal{K} := (2\pi/L)\mathbb{Z}^3$. The modes shall be orthogonal, i.e.

$$\int_{\mathcal{D}} \mathbf{h}_{s_p}(\mathbf{p}) \exp(i\mathbf{p} \cdot \mathbf{x}) \cdot [\mathbf{h}_{s_q}(\mathbf{q}) \exp(i\mathbf{q} \cdot \mathbf{x})]^* \, d\mathbf{x} = C \delta_{\mathbf{p},\mathbf{q}} \delta_{s_p,s_q} \; ; \; C > 0 \text{ const.} \quad (2)$$

The complex vector amplitudes $\mathbf{h}_\pm(\mathbf{k})$ should satisfy $\mathbf{k} \cdot \mathbf{h}_\pm(\mathbf{k}) = 0$ and $i\mathbf{k} \times \mathbf{h}_\pm(\mathbf{k}) = \pm|\mathbf{k}|\mathbf{h}_\pm(\mathbf{k})$. A convenient choice of basis for the $\mathbf{h}_\pm(\mathbf{k})$ is then given by

$$\mathbf{h}_\pm(\mathbf{k}) := \boldsymbol{\nu} \times \boldsymbol{\kappa} \pm i\boldsymbol{\nu}, \text{ with } \boldsymbol{\kappa} := \mathbf{k}/k, \; \boldsymbol{\nu} := \mathbf{k} \times \boldsymbol{\Gamma}/|\mathbf{k} \times \boldsymbol{\Gamma}|, \; \boldsymbol{\Gamma} := \text{ const,} \quad (3)$$

for which $|\mathbf{h}_\pm(\mathbf{k})|^2 := \mathbf{h}_\pm(\mathbf{k}) \cdot \mathbf{h}_\pm(\mathbf{k})^* = 2$ and $\mathbf{h}_\pm(\mathbf{k}) \cdot \mathbf{h}_\mp(\mathbf{k})^* = 0$.

At this point, one notices the key features of the helical modes $\mathbf{h}_\pm(\mathbf{k}) \exp(i\mathbf{k} \cdot \mathbf{x})$ which greatly simplifies analysis of modal expansions of the 3D Euler and related equations, such as 3D Navier-Stokes. Namely, the helical modes $\mathbf{h}_\pm(\mathbf{k}) \exp(i\mathbf{k} \cdot \mathbf{x})$, are divergence-free eigenfunctions of the curl operator. Specifically,

$$\nabla \cdot \mathbf{h}_s(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{x}} = i\mathbf{k} \cdot \mathbf{h}_s(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{x}} = 0 \quad (4)$$

and

$$\nabla \times \mathbf{h}_s(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{x}} = i\mathbf{k} \times \mathbf{h}_s(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{x}} = s|\mathbf{k}|\mathbf{h}_s(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{x}}. \quad (5)$$

See [9, 10, 28, 29] for more information about how this Galerkin decomposition into divergence-free eigenfunctions of the curl are used as a standard tool in analysis of 3D solution behaviour of the deterministic Euler fluid equations and Navier-Stokes fluid equations. In particular, the helical mode expansions in Eqs. (6) comprise the source of the popular *shell models* as finite-dimensional expansions of turbulent fluid dynamics. Thus, this expansion provides a useful framework for studying low-dimension stochastic models of 3D Navier-Stokes turbulence.

In terms of the basis of helical modes $\mathbf{h}_\pm(\mathbf{k}) \exp(i\mathbf{k} \cdot \mathbf{x})$, the divergence-free fluid velocity $\mathbf{v}(\mathbf{x}, t)$ and vorticity $\boldsymbol{\omega}(\mathbf{x}, t)$ are expressed in [28, 29] in terms of complex vector amplitudes $\mathbf{u}(\mathbf{k}, t), \boldsymbol{\varpi}(\mathbf{k}, t) \in \mathbb{C}^3$, respectively,

$$\mathbf{v}(\mathbf{x}, t) := \sum_{\mathbf{p}} \mathbf{u}(\mathbf{p}, t)e^{i\mathbf{p}\cdot\mathbf{x}} := \sum_{\mathbf{p}} \sum_{s_p=\pm} a_{s_p}(\mathbf{p}, t)\mathbf{h}_{s_p}(\mathbf{p})e^{i\mathbf{p}\cdot\mathbf{x}},$$

$$\boldsymbol{\omega}(\mathbf{x}, t) := \sum_{\mathbf{q}} \boldsymbol{\varpi}(\mathbf{q}, t)e^{i\mathbf{q}\cdot\mathbf{x}} := \sum_{\mathbf{q}} \sum_{s_q=\pm} s_q|\mathbf{q}| \, a_{s_q}(\mathbf{q}, t)\mathbf{h}_{s_q}(\mathbf{q})e^{i\mathbf{q}\cdot\mathbf{x}}. \quad (6)$$

Here, the choice

$$\mathbf{u}(\mathbf{k}, t) := a_+(\mathbf{k}, t)\mathbf{h}_+(\mathbf{k}) + a_-(\mathbf{k}, t)\mathbf{h}_-(\mathbf{k}) = \sum_{s_k=\pm} a_{s_k}(\mathbf{k}, t)\mathbf{h}_{s_k}(\mathbf{k}), \tag{7}$$

with $a_s^*(\mathbf{k}) = a_s(-\mathbf{k})$ [29], was made, so that (5) implies

$$\boldsymbol{\varpi}(\mathbf{k}, t) := |\mathbf{k}|\Big(a_+(\mathbf{k}, t)\mathbf{h}_+(\mathbf{k}) - a_-(\mathbf{k}, t)\mathbf{h}_-(\mathbf{k})\Big) = |\mathbf{k}|\sum_{s_k=\pm} s_k\, a_{s_k}(\mathbf{k}, t)\mathbf{h}_{s_k}(\mathbf{k}). \tag{8}$$

The conservation laws for the Euler fluid kinetic energy and helicity—expressed as integrals over the spatially periodic box $\mathcal{D}$—can be evaluated in Fourier space via Parseval's theorem, as follows,

$$\frac{1}{2}\int_{\mathcal{D}} |\mathbf{v}(\mathbf{x}, t)|^2\, d^3x = \frac{1}{2}\sum_{\mathbf{k}} \mathbf{u}(\mathbf{k}, t)\cdot\mathbf{u}^*(\mathbf{k}, t) = \sum_{\mathbf{k}}\sum_{s_k=\pm} a_{s_k}(\mathbf{k}, t)\cdot a_{s_k}^*(\mathbf{k}, t),$$

$$\int_{\mathcal{D}} \mathbf{v}(\mathbf{x}, t)\cdot\mathrm{curl}\mathbf{v}(\mathbf{x}, t)\, d^3x = \sum_{\mathbf{k}} \mathbf{u}(\mathbf{k}, t)\cdot\boldsymbol{\varpi}^*(\mathbf{k}, t)$$

$$= \sum_{\mathbf{k}}\sum_{s_k=\pm} ks_k\, a_{s_k}(\mathbf{k}, t)a_{s_k}^*(\mathbf{k}, t)\, \mathbf{h}_{s_k}(\mathbf{k})\cdot\mathbf{h}_{s_k}^*(\mathbf{k})$$

$$= 2\sum_{\mathbf{k}}\sum_{s_k=\pm} ks_k\, a_{s_k}(\mathbf{k}, t)a_{s_k}^*(\mathbf{k}, t)\,. \tag{9}$$

Expanding the terms of the Euler equation in curl form (57), we obtain the Euler equations for the coefficients $a_{s_k}(\mathbf{k}, t)$. For all $\mathbf{k} \in \mathcal{K}$, $s_k \in \{+, -\}$ we have

$$\partial_t a_{s_k}(\mathbf{k}, t) = -\frac{1}{4}\sum_{\mathbf{p}+\mathbf{q}+\mathbf{k}=0}\sum_{s_p, s_q}(s_p|\mathbf{p}| - s_q|\mathbf{q}|)a_{s_p}^*(\mathbf{p}, t)a_{s_q}^*(\mathbf{q}, t)\mathbf{h}_{s_p}^*(\mathbf{p})\times\mathbf{h}_{s_q}^*(\mathbf{q})\cdot\mathbf{h}_{s_k}^*(\mathbf{k}). \tag{10}$$

For the explicit derivation of Eq. (10) see section "Deterministic Euler" in Appendix 2.

The elementary interactions in Fourier space take place between triads of wave vectors such that $\mathbf{k} + \mathbf{p} + \mathbf{q} = 0$, as exemplified in Eq. (10) above. There are two degrees of freedom per wave vector, (a_+, a_-), so eight different types of interaction are allowed according to the value of the triplet $(s_k, s_p, s_q) = (\pm 1, \pm 1, \pm 1)$. Consider a fixed triple of wave vectors $\mathbf{k}, \mathbf{p}, \mathbf{q} \in \mathcal{K}$ such that $\mathbf{k} + \mathbf{p} + \mathbf{q} = 0$ and a fixed triple $s_k, s_p, s_q \in \{+, -\}$. This gives rise to three coefficients $a_{s_k}(\mathbf{k}, t), a_{s_p}(\mathbf{p}, t), a_{s_q}(\mathbf{q}, t)$, which we compactly summarise into the complex vector

$$\mathbf{a} = (a_{s_k}, a_{s_p}, a_{s_q}) \in \mathbb{C}^3 .$$

The dynamics of $\mathbf{a}$ is determined by the three equations obtained from (10)

$$\frac{da_{s_k}}{dt} = g(s_p|\mathbf{p}| - s_q|\mathbf{q}|)a_{s_p}^* a_{s_q}^* + R, \tag{11}$$

$$\frac{da_{s_p}}{dt} = g(s_q|\mathbf{q}| - s_k|\mathbf{k}|)a_{s_q}^* a_{s_k}^* + R, \tag{12}$$

$$\frac{da_{s_q}}{dt} = g(s_k|\mathbf{k}| - s_p|\mathbf{p}|)a_{s_k}^* a_{s_p}^* + R, \tag{13}$$

We pick out a summand and cycle through k, p, q

$$\frac{da_{s_k}}{dt} = g(s_p|\mathbf{p}| - s_q|\mathbf{q}|)a_{s_p}^* a_{s_q}^*, \tag{14}$$

$$\frac{da_{s_p}}{dt} = g(s_q|\mathbf{q}| - s_k|\mathbf{k}|)a_{s_q}^* a_{s_k}^*, \tag{15}$$

$$\frac{da_{s_q}}{dt} = g(s_k|\mathbf{k}| - s_p|\mathbf{p}|)a_{s_k}^* a_{s_p}^*, \tag{16}$$

with the constant complex scalar

$$g := -\frac{1}{4}\mathbf{h}_{s_p}^*(\mathbf{p}) \times \mathbf{h}_{s_q}^*(\mathbf{q}) \cdot \mathbf{h}_{s_k}^*(\mathbf{k}). \tag{17}$$

The equations corresponding to the single triad interaction of type (s_k, s_p, s_q) with $\mathbf{k} + \mathbf{p} + \mathbf{q} = 0$ thus have the complex vector form also derived in [28],

$$\frac{d\mathbf{a}}{dt} = g\mathbf{a}^* \times \mathbb{D}\mathbf{a}^* = g(\mathbf{a} \times \mathbb{D}\mathbf{a})^*, \tag{18}$$

with the constant diagonal matrix

$$\mathbb{D} := \mathrm{diag}\big(s_k|\mathbf{k}|, s_p|\mathbf{p}|, s_q|\mathbf{q}|\big). \tag{19}$$

The form of the factor g defined in (17) above can be calculated from (3) to show that it depends on the shape and the orientation of the wave-vector triad, but not on its scale; since the real and imaginary parts of the complex helical vector amplitudes $\mathbf{h}_{s_k}$ are unit vectors. Moreover, $\mathbb{D}\mathbf{a}$ can be seen to represent the (s_k, s_p, s_q) components of the vorticity vector amplitude ϖ through Eq. (8) above. Two conservation laws for real-valued triad energy and helicity follow immediately from Eq. (18), as

$$\frac{d}{dt}(\mathbf{a} \cdot \mathbf{a}^*) = 0 \quad \text{and} \quad \frac{d}{dt}(\mathbf{a} \cdot \mathbb{D}\mathbf{a}^*) = 0 . \tag{20}$$

The dynamical system in Eq. (18) is similar to rigid body dynamics, but replaced by complex angular momentum (**a**), complex angular velocity ($\mathbb{D}$**a**) and real moment of inertia $\mathbb{I} = \mathbb{D}^{-1}$ relating the two complex quantities.[1]

2.2 Stochastic Parametrizations for the 3D Euler Equation

In the following we introduce a reduced order model for the stochastic parametrizations introduced through the Stochastic Advection by Lie Transport paradigm as well as the Location Uncertainty paradigm. We explain below the rationale of these parametrizations:

2.2.1 Modelling Under the Stochastic Advection by Lie Transport Principle

The SALT equations were derived in [16] using a Stratonovich stochastic version of Hamilton's variational principle [18] in combination with Kraichnan's scalar turbulence model based on Stratonovich stochastic Lagrangian paths [20]. The application of Hamilton's principle with an imposed stochastic Lie transport constraint implied an Euler-Poincaré equation for the fluid motion [18]. The 3D SALT Euler equations for divergence-free fluid velocity $\mathbf{v}(\mathbf{x}, t)$ are given by

$$d\mathbf{v} + (\,d\mathbf{x}_t \cdot \nabla)\mathbf{v} + v_j \nabla \, d\mathbf{x}_t^j = -\nabla \, dp \,,$$
$$\text{with} \quad d\mathbf{x}_t = \mathbf{v} dt + \sum_i \boldsymbol{\xi}_i(\mathbf{x}) \circ dW_t^i \,. \tag{21}$$

As discussed in [16], this motion equation yields a Kelvin-Noether circulation theorem for the stochastic system

$$d \oint_{c(\mathbf{x}_t)} \mathbf{v} \cdot d\mathbf{x} = -\oint_{c(\mathbf{x}_t)} \nabla \, dp \cdot d\mathbf{x} = 0 \,. \tag{22}$$

This stochastic Kelvin circulation theorem is has the same form as that for the deterministic system, except that each line element of the material loop in Kelvin's theorem follows the Stratonovich stochastic Lagrangian path, $\mathbf{x}_t$.

The real vectors $\boldsymbol{\xi}_i$ comprise the time-independent noise amplitudes which are to be determined from data assimilation. The W^i are independent (uni-dimensional) standard Brownian motions and $\circ$ denotes stochastic integration in the Stratonovich sense.[2] The curl form of the SALT Euler motion equation in (21) is obtained

---

[1] Rigid body dynamics with complex angular momentum has also been discussed previously in [3].

[2] An exposition of Brownian motion, stochastic calculus and the Stratonovich integral is to be found, for example, in the monograph [19].

from (56) and given by

$$d\mathbf{v} - d\mathbf{x}_t \times \text{curl}\mathbf{v} + \nabla(\mathbf{v} \cdot d\mathbf{x}_t) = -\nabla \, dp. \tag{23}$$

The motion equation (23) and its curl yielding the SALT vorticity equation implies a formula for the evolution of the helicity of the flow, Λ, defined as

$$\Lambda := \int_{\mathcal{D}} \mathbf{v} \cdot \text{curl}\mathbf{v} \, d^3 x \,. \tag{24}$$

Upon applying the divergence theorem, one finds

$$d\Lambda = - \int_{\partial \mathcal{D}} \widehat{\mathbf{n}} \cdot \left((\mathbf{v} \cdot \text{curl}\mathbf{v}) \, d\mathbf{x}_t + \text{curl}\mathbf{v} \, dp \right) dS \,. \tag{25}$$

Thus, in a periodic 3D domain, or in an infinite 3D domain with asymptotically vanishing boundary conditions, the SALT motion equation in (21) or (23) preserves the helicity, Λ, defined in (24). However, a glance at the SALT motion equation in (23) informs us that it will not preserve the kinetic energy, since even with the usual fluid boundary conditions $\text{div}\mathbf{v} = 0$ implies

$$\tfrac{1}{2} \, d\|\mathbf{v}\|_{L^2}^2 := \int_{\mathcal{D}} \mathbf{v} \cdot d\mathbf{x}_t \times \text{curl}\mathbf{v} \, d^3 x \neq 0 \,. \tag{26}$$

2.2.2   Modeling Under the Location Uncertainty Principle

The Location Uncertainty principle consists in decomposing the flow trajectory $\mathbf{x} \colon \Omega \times \mathbb{R}^+ \to \Omega$ over a bounded domain, $\Omega \subset \mathbb{R}^3$

$$d\mathbf{x}_t = \mathbf{v}\,(\mathbf{x}_t, t) \, dt + \boldsymbol{\sigma}\,(\mathbf{x}_t, t) \, d\mathbf{W}_t \tag{27}$$

in terms of $\mathbf{v}\,(\mathbf{x}_t, t)$, a smooth-in-time component of the (Lagrangian) velocity and a noise $\boldsymbol{\sigma}\,(\mathbf{x}_t, t) \, d\mathbf{W}_t$, which has here to be understood in the Itô sense and that accounts for the unresolved processes. The Wiener process, $\mathbf{W}_t$ is a H-valued (cylindrical) Brownian motion, where H is the Hilbert space of square integrable functions. The noise is then properly defined as the application of an Hilbert-Schmidt symmetric integral kernel $\boldsymbol{\sigma}_t \boldsymbol{f}\,(\boldsymbol{x}) = \int_{\mathcal{S}} \breve{\sigma}\,(\boldsymbol{x}, \boldsymbol{y}, t)\, \boldsymbol{f}\,(\boldsymbol{y}) \, d\boldsymbol{y}$ to the H-valued cylindrical Wiener process $\mathbf{W}$ as

$$(\boldsymbol{\sigma}_t d\mathbf{W}_t)^i \,(\boldsymbol{x}) = \int_{\mathcal{S}} \breve{\sigma}_{ik}\,(\boldsymbol{x}, \boldsymbol{y}, t) \, dW_t^k\,(\boldsymbol{y}) \, d\boldsymbol{y}, \tag{28}$$

The role of the integrable kernel $\breve{\sigma}$ is to impose a spatial correlation on the small-scale component. It leads to the covariance tensor $\boldsymbol{Q}$

$$Q_{ij}(\boldsymbol{x}, \boldsymbol{y}, t, s) = \mathbb{E}\left[(\boldsymbol{\sigma}_t \mathrm{d}\mathbf{W}_t(\boldsymbol{x}))^i (\boldsymbol{\sigma}_t \mathrm{d}\mathbf{W}_s(\boldsymbol{y}))^j\right]$$

$$= \delta(t - s)\,\mathrm{d}t \int_{\mathcal{S}} \breve{\sigma}_{ik}(\boldsymbol{x}, \boldsymbol{z}, t)\,\breve{\sigma}_{kj}(\boldsymbol{z}, \boldsymbol{y}, s)\,\mathrm{d}\boldsymbol{z},$$

of the centered Gaussian process $\boldsymbol{\sigma}_t \mathrm{d}\mathbf{W}_t \sim \mathcal{N}(0, \mathbf{Q}\mathrm{d}t)$. The diagonal components of the covariance tensor per unit of time, referred to as the variance tensor, $\mathbf{a}$, is a positive definite matrix defined as $\mathbf{a}(\mathbf{x}, t)\delta(t - t')\mathrm{d}t = \mathbf{Q}(\mathbf{x}, \mathbf{x}, t, t')$, that quantifies the strength of the noise and has the dimension of a viscosity in $\mathrm{m}^2\mathrm{s}^{-1}$. The operator Q being compact auto-adjoint positive definite operator on H, it admits eigenfunctions $\boldsymbol{\xi}_n(\cdot, t)$ with (strictly) positive eigenvalues $\lambda_n(t)$ satisfying $\sum_{n\in\mathbb{N}} \lambda_n(t) < +\infty$. As a consequence, the noise and the variance tensor $\boldsymbol{a}$ can be expressed through the spectral representation

$$\boldsymbol{\sigma}_t \mathrm{d}\mathbf{W}_t(\boldsymbol{x}) = \sum_{n\in\mathbb{N}} \lambda_n^{1/2}(t)\,\boldsymbol{\xi}_n(\boldsymbol{x}, t)\,\mathrm{d}\beta_n \tag{29}$$

$$\boldsymbol{a}(\boldsymbol{x}, t) = \sum_{n\in\mathbb{N}} \lambda_n(t)\,\boldsymbol{\xi}_n(\boldsymbol{x}, t)\,\boldsymbol{\xi}_n^{\dagger}(\boldsymbol{x}, t). \tag{30}$$

The rate of change of a volume V_t of the scalar q is given by the stochastic Reynolds transport theorem, introduced in [23]

$$\mathrm{d}\int_{V_t} q(\boldsymbol{x}, t)\,\mathrm{d}\boldsymbol{x} = \int_{V_t} \left\{\mathrm{D}_t q + q\nabla\cdot[\mathbf{v}^{\star}\,\mathrm{d}t + \boldsymbol{\sigma}_t \mathrm{d}\mathbf{W}_t]\right\}(\boldsymbol{x}, t)\,\mathrm{d}\boldsymbol{x}, \tag{31}$$

with the transport operator

$$\mathrm{D}_t q = \mathrm{d}_t q + \left[\mathbf{v}^{\star}\,\mathrm{d}t + \boldsymbol{\sigma}_t\,\mathrm{d}\mathbf{W}_t\right]\cdot\nabla q - \frac{1}{2}\nabla\cdot(\mathbf{a}\nabla q)\,\mathrm{d}t. \tag{32}$$

In this formula, the first component of the right-hand side is the increment in time at a fixed location of the process q, that is $\mathrm{d}_t q = q(\mathbf{x}_t, t + \mathrm{d}t) - q(\mathbf{x}_t, t)$, playing the role of a derivative in time for a non differentiable process. The effective velocity $\mathbf{v}^{\star}$ is defined as

$$\mathbf{v}^{\star} = \mathbf{v} - \frac{1}{2}\nabla\cdot\mathbf{a} + \boldsymbol{\sigma}_t^{*}(\nabla\cdot\boldsymbol{\sigma}_t), \tag{33}$$

where the velocity component $\mathbf{v}_s = \frac{1}{2}\nabla\cdot\mathbf{a}$ results from the noise inhomogeneities. For incompressible homogeneous noise as considered in this work $\mathbf{v}^{\star} = \mathbf{v}$. Besides, the diffusion term exactly balances the noise brought by the noise. With Stratonovich convention and a homogeneous noise the transport operator takes a

simplified form similar to the material derivative:

$$D_t q = d_t q + (\mathbf{v}\,dt + \boldsymbol{\sigma}_t \circ d\mathbf{W}_t) \cdot \nabla q. \tag{34}$$

For a divergence free homogeneous noise, the Euler equation, in the LU framework, can then be defined as:

$$d_t \mathbf{v} + (\mathbf{v}\,dt + \boldsymbol{\sigma}_t \circ d\mathbf{W}_t) \cdot \nabla \mathbf{v} = -\nabla dp_t, \quad \nabla \cdot \mathbf{v} = 0, \tag{35}$$

where dp_t denotes the pressure composed of a finite variation term and a martingale pressure term. With the Leray projection, $\mathbb{P}$, this pressure term can be removed and we obtain, the inertial form of the Euler equation:

$$d_t \mathbf{v} + \mathbb{P}\big((d\mathbf{x}_t \cdot \nabla)\mathbf{v}\big) = 0, \quad \nabla \cdot \mathbf{v} = 0. \tag{36}$$

2.3 Triad Model Comparison

The reduced order model for the full-scale 3D SALT Euler and 3D LU Euler for a single triad interaction equation is obtained by projecting the continuous stochastic Euler models onto the helical modes, in the same fashion as for the deterministic equation (18). Therefore, we introduce an additional Stratonovich stochastic term into the transport velocity in (7) as

$$d\mathbf{x}_t(\mathbf{k}, t) := \big(a_+(\mathbf{k}, t)\mathbf{h}_+ + a_-(\mathbf{k}, t)\mathbf{h}_-\big)\,dt + \sum_i \big(b_+^i(\mathbf{k})\mathbf{h}_+ + b_-^i(\mathbf{k})\mathbf{h}_-\big) \circ dW_t^i, \tag{37}$$

where the $\mathbf{k}$-dependent complex vector $\mathbf{b}(\mathbf{k}) := (b_{s_k}, b_{s_p}, b_{s_q})^T \in \mathbb{C}^3$ represents the time-independent noise amplitude which is to be determined from data assimilation, similar to the continuous stochastic models (21). Enumerating the equation for a single triad then yields, after rearranging using exchange symmetry in $(\mathbf{k}, \mathbf{p}, \mathbf{q})$, the matrix equation

$$d \begin{bmatrix} a_{s_k} \\ a_{s_p} \\ a_{s_q} \end{bmatrix} = g \begin{bmatrix} 0 & -qs_q a_{s_q} & ps_p a_{s_p} \\ qs_q a_{s_q} & 0 & -ks_k a_{s_k} \\ -ps_p a_{s_p} & ks_k a_{s_k} & 0 \end{bmatrix}^* \begin{bmatrix} a_{s_k}\,dt + b_{s_k} \circ dW_t \\ a_{s_p}\,dt + b_{s_p} \circ dW_t \\ a_{s_q}\,dt + b_{s_q} \circ dW_t \end{bmatrix}^* . \tag{38}$$

Upon applying the previous steps for the deterministic case to the stochastic velocity in (37), the single triad interaction dynamics for the SALT case would emerge as, cf. Eq. (18),

$$d\mathbf{a} = g\Big(\mathbf{a}(\mathbf{k}, t)dt + \mathbf{b}(\mathbf{k}) \circ dW_t\Big)^* \times \mathbb{D}\mathbf{a}^* . \tag{39}$$

The details of this computation can be found in section "LU Euler" in Appendix 2 for the 3D LU Euler model and in section "SALT Euler" in Appendix 2 for the 3D SALT Euler model.

Remarkably, the HST equation for triad interaction (39) still preserves the triad helicity $\mathbf{a} \cdot \mathbb{D}\mathbf{a}^*$. Hence we name this model the *helicity preserving stochastic triad* (HST) model. Note that in both equations we use as single source of noise (One Brownian motion drives the entire system).

It is readily checked that the HST triad evolution (39) preserves the helicity. Let's have a look at the diffusion coefficients.

$$\mathbf{a}^* \times \mathbb{D}\mathbf{b} = \begin{bmatrix} a_p^* s_q q b_q - a_q^* s_p p b_p \\ a_q^* s_k k b_k - a_k^* s_q q b_q \\ a_k^* s_p p b_p - a_p^* s_k k b_k \end{bmatrix}, \quad \mathbf{b} \times \mathbb{D}\mathbf{a}^* = \begin{bmatrix} b_p s_q q a_q^* - b_q s_p p a_p^* \\ b_q s_k k a_k^* - b_k s_q q a_q^* \\ b_k s_p p a_p^* - b_p s_k k a_k^* \end{bmatrix}. \tag{40}$$

Taking the difference

$$\mathbf{a}^* \times \mathbb{D}\mathbf{b} - \mathbf{b} \times \mathbb{D}\mathbf{a}^* = \begin{bmatrix} a_p^*(s_q q + s_p p)b_q - a_q^*(s_p p + s_q q)b_p \\ a_q^*(s_k k + s_q q)b_k - a_k^*(s_q q + s_k k)b_q \\ a_k^*(s_p p + s_k k)b_p - a_p^*(s_k k + s_p p)b_k \end{bmatrix}. \tag{41}$$

Writing $\rho := \operatorname{Tr}\mathbb{D}$ we get

$$\mathbf{a}^* \times \mathbb{D}\mathbf{b} - \mathbf{b} \times \mathbb{D}\mathbf{a}^* = \begin{bmatrix} a_p^*(\rho - s_k k)b_q - a_q^*(\rho - s_k k)b_p \\ a_q^*(\rho - s_p p)b_k - a_k^*(\rho - s_p p)b_q \\ a_k^*(\rho - s_q q)b_p - a_p^*(\rho - s_q q)b_k \end{bmatrix}. \tag{42}$$

So that the difference term becomes

$$\mathbf{a}^* \times \mathbb{D}\mathbf{b} - \mathbf{b} \times \mathbb{D}\mathbf{a}^* = (\rho \operatorname{Id} - \mathbb{D})(\mathbf{a} \times \mathbf{b})^*. \tag{43}$$

Since the projections of the LU and SALT models onto a single triad are indistinguishable, we introduce a different model that conserves energy on a single triad to enable a comparison between energy conserving and helicity conserving models.

Energy-Preserving Stochastic Triad (EST) Model We introduce below a modified version of the HST triad equation (39) that introduces stochasticity into the vorticity instead of into the transport velocity and thereby conserves the energy. This is inspired by the full-scale model introduced in [17]. The reduced model is as follows

$$d\mathbf{a} = -g\mathbf{a}^* \times \mathbb{D}\Big(\mathbf{a}(\mathbf{k}, t)\, dt + \mathbf{b}(\mathbf{k}) \circ dW_t\Big)^*. \tag{44}$$

We call this model the energy preserving stochastic triad (EST). Written in matrix form Eq. (44) becomes

$$
\mathrm{d}
\begin{bmatrix}
a_{s_k} \\
a_{s_p} \\
a_{s_q}
\end{bmatrix}
= g
\begin{bmatrix}
0 & -a_{s_q} & a_{s_p} \\
a_{s_q} & 0 & -a_{s_k} \\
-a_{s_p} & a_{s_k} & 0
\end{bmatrix}^*
\begin{bmatrix}
k a_k \big(a_{s_k}\,\mathrm{d}t \,+\, b_{s_k}(\mathbf{k}) \circ dW_t \big) \\
p s_p \big(a_{s_p}\,\mathrm{d}t \,+\, b_{s_p}(\mathbf{p}) \circ dW_t \big) \\
q s_q \big(a_{s_q}\,\mathrm{d}t \,+\, b_{s_q}(\mathbf{q}) \circ dW_t \big)
\end{bmatrix}^*.
\tag{45}
$$

The exchange symmetry between the two models HST and EST in the placement of the noise in Eqs. (39) and (44) is apparent already in the exchange symmetry between velocity and vorticity in Euler's fluid equations (1).

Deviation from the Conservation Laws We can write the equations for the deviation from the conservation laws, which is present in both models. The SALT model deviates from the energy conservation by

$$
\mathrm{d}_t E_{\mathrm{HST}} = g\mathbf{b} \cdot (\mathbb{D}\mathbf{a}^* \times \mathbf{a}^*) \circ \mathrm{d}W_t
\tag{46}
$$

whereas the LU model deviates from the helicity conservation by

$$
\mathrm{d}_t H_{\mathrm{EST}} = g\mathbb{D}\mathbf{b} \cdot (\mathbb{D}\mathbf{a}^* \times \mathbf{a}^*) \circ \mathrm{d}W_t.
\tag{47}
$$

This is seen, since, to get the energy we dot the HST equation with $\mathbf{a}^*$ and to get helicity we dot the EST equation with $\mathbb{D}\mathbf{a}^*$ and use the standard identities

$$
\mathbf{a}^* \cdot (\mathbf{b} \times \mathbb{D}\mathbf{a}^*) = \mathbf{b} \cdot (\mathbb{D}\mathbf{a}^* \times \mathbf{a}^*)
\tag{48}
$$

$$
\mathbb{D}\mathbf{a}^* \cdot (\mathbf{a}^* \times \mathbb{D}\mathbf{b}) = \mathbb{D}\mathbf{b} \cdot (\mathbb{D}\mathbf{a}^* \times \mathbf{a}^*).
\tag{49}
$$

Therefore, $\mathbf{b}$ respects the right scaling and no further scale adjustments between the SALT and LU noise scaling need to be performed in order to compare the models.

3 Data Assimilation Comparison

In this section, we perform a comparative study of the two reduced order models (HST and EST) introduced above by using data assimilation tools. The particular methodology that we make use of is that of particle filters. We will first briefly explain the particle filtering methodology in a generic framework:

Let X and Z be two processes defined on a given probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The process X is usually called the *signal process* or the *truth* and Z is the *observation process*. In this paper, X is the pathwise solution of a (deterministic) shell model. The pair of processes (X, Z) forms the basis of the nonlinear filtering problem which consists in finding the best approximation of the posterior

distribution of the signal X_t given the observations $Z_1, Z_2, \ldots, Z_{t_n}$.[3] The posterior distribution of the signal at time t is denoted by π_t. We let d_X be the dimension of the state space and d_Z be the dimension of the observation space. This mixed continuous-discrete time framework can be embedded into a fully discrete framework, whereby one is interested in computing the conditional probability law of the signal at the time corresponding to the observation time. in other words one wants to compute the conditional distribution π_{t_n} of $X(t_n)$ given the data $Z(t_1), Z(t_2), \ldots, Z(t_n)$. The process X is assumed to be a Markov process, and we will denote by $\mathcal{K}_n$ its transition kernel, that is

$$\mathcal{K}_n : \mathbb{R}^{d_X} \times \mathcal{B}(\mathbb{R}^{d_X}) \to [0, 1], \ \mathcal{K}_n(x, B) = \mathbb{P}(X_{t_n} \in B | X_{t_{n-1}} = x) \tag{50}$$

for any Borel measurable set $B \in \mathcal{B}(\mathbb{R}^{d_X})$ and $x \in \mathbb{R}^{d_X}$. The process Z models noisy measurements of the truth, using the so-called *observation operator* $\mathscr{H}$: $\mathbb{R}^{d_X} \to \mathbb{R}^{d_Z}$:

$$Z_n = \mathscr{H}(X_{t_n}) + V_n \tag{51}$$

where $\{V_n\}_{n \geq 0}$ are independent identically distributed random variables that represent the measurement noise and $\mathscr{H}$ is a Borel-measurable function. In this paper we will assume that $\{V_n\}_{n \geq 0}$ have standard normal distributions, but the same methodology can be applied to more general distributions. Observations are incorporated into the system at *assimilation times*. The following recursion formula holds (see [2])

$$\pi_n = g_n \star \pi_{n-1} \mathcal{K}_n \tag{52}$$

where by '$\star$' we denoted the projective product (see e.g. Definition 10.4 in [2]).

In the following, we compare approximations of the posterior distribution of the signal using *particle filters*. These are sequential Monte Carlo methods which generate approximations of the posterior distribution π_t using sets of *particles*. That is, they generate approximations that are (random) measures of the form

$$\pi_n \approx \sum_\ell \mathrm{w}_n^\ell \delta(x_n^\ell),$$

where δ is the Dirac delta distribution, $\mathrm{w}_t^1, \mathrm{w}_t^2, \ldots$ are the *weights* of the particles and $x_t^1, x_t^2, \ldots$ are their corresponding positions. Particle filters are used to make inferences about the signal process by using Bayes' theorem, the time-evolution induced by the signal X, and the observation process Z.

---

[3] For a mathematical introduction on the subject, see e.g. [2].

In a standard particle filter, the particles evolve between assimilation times according to the law of the signal. As we explain below, at each assimilation time the observation is incorporated into the system through the *likelihood function*:

$$g_t^{z_t} : \mathbb{R}^{d_X} \to \mathbb{R}_+, \ g_t^{z_t}(x) = g_t(z_t - \mathcal{H}(x)) \ \text{ such that } \ \mathbb{P}(Z_t \in dz_t | X_t = x) = g_t^{z_t}(x) dz_t \tag{53}$$

and all particles are weighted depending on the likelihood of their corresponding position, given the observation. More precisely, the particle ℓ is given the weight $w_n^l = g_n^{Z_n}(x_\ell)$. Heuristically, the particle weight measures how close the particle trajectory is to the signal trajectory. A selection procedure is then applied to the set of weighted particles. Particles with higher conditional likelihood (guided by the observation) have higher weights and will be multiplied, while those which have small likelihoods will be eliminated. For the basic particle filter, this is done by sampling with replacement from the population of particles, with corresponding probabilities proportional to their weights.

A Monte Carlo implementation of the transition kernel of the signal may not always yield good approximations. In many situations one replaces the original transition kernel with likelihood informed importance proposals, leading to much better approximations. One situation when this is necessary is when the original process is actually deterministic (aside for the initial position which is assumed to be random). This is the case in our paper.

To overcome the collapse of the particle filter when using deterministic transition kernels, one can use a Markov Chain Monte Carlo procedure that leaves the deterministic dynamics invariant. This procedure can be costly and might not always introduce enough spread into the sample. In this paper, we propose a different approach, which we illustrate numerically in Sect. 3.1.2 below. In particular, we propose two different transition kernels based on the physical conservation properties:

- The transition kernel associated with the HST model equation (38). We will denote this transition kernel by $\mathcal{K}^{\text{HST}}$. As we have explained above, this transition kernel preserves the helicity of the system.
- The transition kernel associated with the triad model equation (45) we will denote this transition kernel by $\mathcal{K}^{\text{EST}}$. As we have explained above, this transition kernel preserves the energy of the system.

3.1 Numerical Studies

3.1.1 Numerical Implementation

The models are discretised using the stochastic SSPRK3 scheme which is documented, for example, in [11]. In our specific case, for instance, the HST model is discretised as

$$\mathbf{q}_1^n = \mathbf{a}_n + g(\mathbf{a}_n^* \times \mathbb{D}\mathbf{a}_n^*)\Delta t + g(\mathbf{b} \times \mathbb{D}\mathbf{a}_n^*)\Delta W$$

$$\mathbf{q}_2^n = (3/4)\mathbf{a}_n + (1/4)(\mathbf{q}_1^n + g((\mathbf{q}_1^n)^* \times \mathbb{D}(\mathbf{q}_1^n)^*)\Delta t + g(\mathbf{b} \times \mathbb{D}(\mathbf{q}_1^n)^*)\Delta W)$$

$$\mathbf{a}_{n+1} = (1/3)\mathbf{a}_n + (2/3)(\mathbf{q}_2^n + g((\mathbf{q}_2^n)^* \times \mathbb{D}(\mathbf{q}_2^n)^*)\Delta t + g(\mathbf{b} \times \mathbb{D}(\mathbf{q}_2^n)^*)\Delta W)$$

where Δt denotes the timestep and ΔW the increment of the driving Brownian motion. Further, $\mathbf{a}_n$ is the approximate complex vector amplitude at time $t = n\Delta t$. The EST model is discretised completely analogously. For the numerical simulations we chose the following triad throughout. We set

$$\mathbf{k} = [1, 0, 0], \quad \mathbf{p} = [0, -1, 1], \quad \mathbf{q} = [-1, 1, -1], \tag{54}$$

with parities $s_k = 1$, $s_p = -1$, $s_q = -1$ and the initial value $\mathbf{a}_0 = \frac{1}{\sqrt{3}}[1, 1, 1]$. We set the parameter $\mathbf{\Gamma} = [1, 1, 1]$ and used a time stepsize of $\Delta t = 0.0005$.

3.1.2 Data Assimilation for the Deterministic Model

We illustrate the failure of the particle filter with deterministic transition kernel in Fig. 1. In this case, the particle filtering is performed for an ensemble of $n = 25$ particles evolving according to the deterministic triad dynamics. The initial ensemble is spread around the initial value $\mathbf{a}_0$ of the signal according to a Gaussian distribution with standard deviation $1/\sqrt{600}$ and, in particular, does not contain the true initial point. Data assimilation is performed every 10 time units and the observations are taken from the modal energies of the truth with an observation error η distributed as $\eta \sim \mathcal{N}(\mathbf{0}, \mathrm{C})$ with covariance $\mathrm{C} = \mathrm{diag}(0.005^2, 0.05^2, 0.05^2)$. We observe that both the bias and the RMSE keep increasing with time to values much larger than the observation error. Moreover, the number of *distinct* particles decreases rapidly: after 30 steps, a single particle remains that is not the true particle since the true particle was not part of the initial cloud. The particle filter does not work.
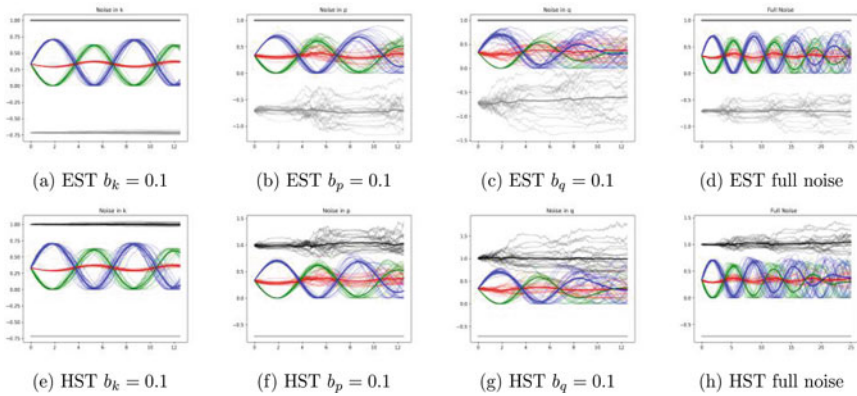
3.1.3 Reduced Order Model Realisations

The deterministic model in Fig. 1a exhibits continually oscillating triad amplitudes. Plotted are the modal energies. Writing $\mathbf{a} = [a_k, a_p, a_q]$ we call the real value $a_k a_k^*$ the *energy of mode k*. Similarly for the two other modes.

We simulate the model realisations for different noise scenarios. We simulate the effect of noise in each single mode. Let the noise amplitude vector be $\mathbf{b} = [b_k, b_p, b_q]$. Then we simulate the two models for $\mathbf{b} = [b_k = 0.1, b_p = 0, b_q = 0]$, $\mathbf{b} = [b_k = 0, b_p = 0.1, b_q = 0]$, and $\mathbf{b} = [b_k = 0, b_p = 0, b_q = 0.1]$. The trajectories of the modal energies for $n = 20$ realisations of the driving noise for each scenario are shown in Fig. 2. We also simulate the case of full noise for

(a) Deterministic Model

(b) Ensemble for Mode **p**

(c) Unique Particles

(d) Bias

(e) RMSE
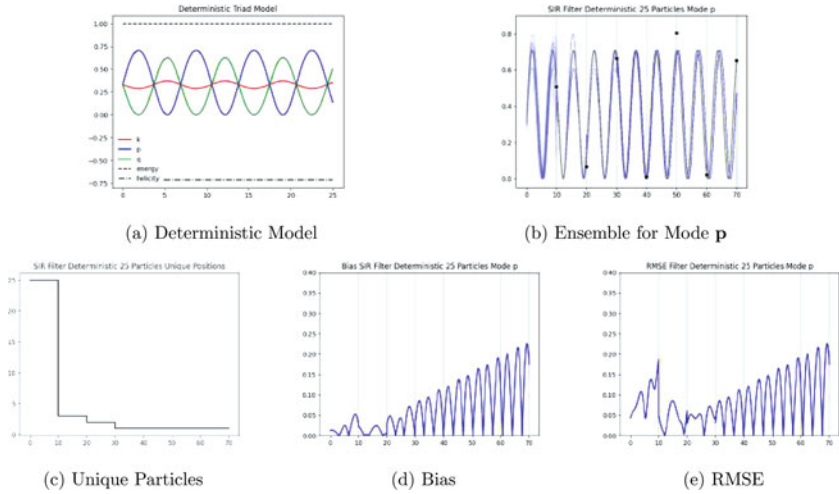
Fig. 1 Deterministic Triad Model. The horizontal axis shows time. (**a**) Evolution of the modal energies (colored lines) as well as the total energy (dashed line) and helicity (dash-dotted line). The deterministic model exhibits continually oscillating modal energies. The simulation also confirms the conservation of energy and helicity. (**b–e**) Data assimilation for the deterministic model using particle filter. (**b**) Evolution of the energy of mode **p** of the signal (grey line) and evolution of the energy of mode **p** for the particle ensemble (blue lines). Noisy observations (black stars) are made and assimilated every 10 time units. (**c**) The number of unique particle positions in the filtering ensemble. (**d**) The bias of the particle ensemble wrt. the observations. (**e**) The RMSE of the particle ensemble wrt. the observations

(a) EST $b_k = 0.1$

(b) EST $b_p = 0.1$

(c) EST $b_q = 0.1$

(d) EST full noise

(e) HST $b_k = 0.1$

(f) HST $b_p = 0.1$

(g) HST $b_q = 0.1$

(h) HST full noise

Fig. 2 Model realisations for both stochastic triad models. Plotted are the modal energies (colored lines), total energy (black line) and helicity (grey line). The respective thick lines are the ensemble means, and the thin lines represent the different stochastic realisations. (**a** + **e**) The noise coefficient **b** = [0.1, 0, 0]. (**b** + **f**) The noise coefficient **b** = [0, 0.1, 0]. (c+g) The noise coefficient **b** = [0, 0, 0.1]. (**d** + **h**) The noise coefficient **b** = [0.1, 0.05, 0.01]
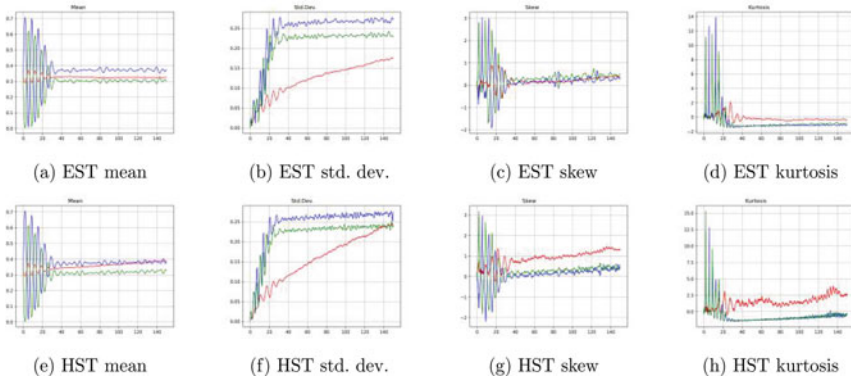
Fig. 3 Evolution of statistical moments of the modal energies for both stochastic models in the full noise case. The statistics are computed pointwise in time from an ensemble of 1000 realisations up to a final time of 150. (**a** + **e**) Ensemble mean. (**b** + **f**) Ensemble standard deviation. (**c** + **g**) Ensemble skew. (**d** + **h**) Ensemble kurtosis

the noise amplitude vector $\mathbf{b} = [b_k = 0.1, b_p = 0.05, b_q = 0.01]$ which was calibrated to the data assimilation objective using the procedure explained in section "Calibration of the Noise Amplitude" in Appendix 3. The ensemble of $n = 20$ realisations of the driving noise in the full noise case is depicted in Fig. 2d and h. In all cases, it can be observed that the mean energy amplitudes of the modes are dampened in both stochastic models. Furthermore, we can experimentally verify the conservation of triad energy for the EST model and the conservation of triad helicity for the HST model.

3.1.4 Model Statistics

Figure 3 shows various statistics of the generated ensembles of $n = 1000$ particles for the HST and EST triad models in the full noise case introduced above. We plot the ensemble mean, standard deviation, skew, and kurtosis.

The effect of large noise coefficients is exemplified in Fig. 4. We observe that the HST model explodes whereas the EST model is more tolerant to large noise coefficients, and even in the extreme case, does not become unstable in the mean. The ensemble means are computed from $n = 500$ realisations, using the noise coefficient $\mathbf{b} = [0.0, 1.0, 0.0]$. Moreover, to stress the EST model, we also ran the same experiment with a noise coefficient of $\mathbf{b} = [0.0, 10.0, 0.0]$ for the EST model alone.

The mean ensemble for a large number of particles, $n = 20{,}000$, is shown in Fig. 5. We can observe that, compared to Fig. 3a and e the oscillations after time 40 are reduced for a very large number of particles. Hence, we believe that the system stabilizes in the mean to stationary modal energies as the limiting effect of the noise.

(a) HST   (b) EST   (c) EST strong noise
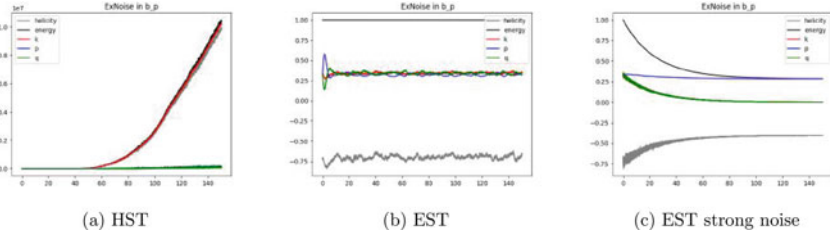
Fig. 4 The effect of large noise coefficients on the mean. Evolution of the mean modal energies (colored lines), mean total energy (black), and mean helicity (grey) for the HST (**a**) and EST (**b**) model with noise coefficient $\mathbf{b} = [0, 1, 0]$. (**c**) Evolution of the mean modal energies, mean total energy, and mean helicity for the EST model with the strong noise coefficient $\mathbf{b} = [0, 10, 0]$

(a) EST   (b) HST
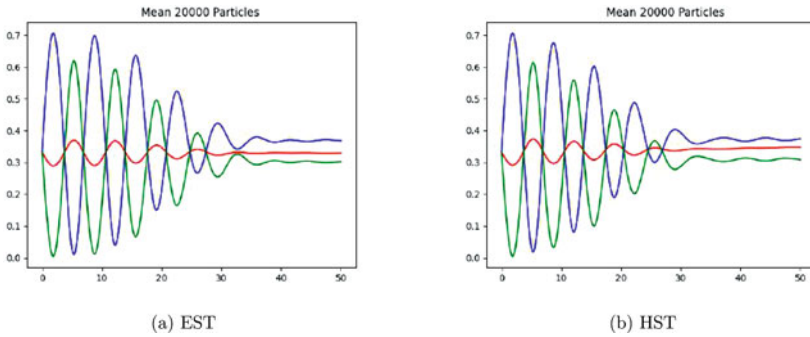
Fig. 5 Evolution of mean modal energies for a very large number of realisations for the EST (**a**) and HST (**b**) models. The mean is computed from 20,000 particles in the full noise case

3.1.5 Data Assimilation

Using the two stochastic models in the full noise case described above, we perform the data assimilation tests using the following framework:

The signal process (the *truth*) is given by the deterministic triad model. The observations are the modal energies of the deterministic model, observed every 10 time units, and perturbed by noise of the form

$$\eta \sim \mathcal{N}(\mathbf{0}, \mathbf{C}), \tag{55}$$

where the covariance matrix $\mathbf{C} \in \mathbb{R}^3$ is chosen to be the diagonal matrix $\mathbf{C} = \text{diag}(0.005^2, 0.05^2, 0.05^2)$.

We use the sequential importance resampling (SIR) particle filter to assimilate the periodically observed signal process under the influence of observation noise. The particles evolve according to the stochastic triad models. Figures 6 and 7 show the results of filtering the ensemble of $n = 100$ particles of the EST and HST triad models, respectively. The ensembles are assessed in terms of the bias and RMSE statistics. We analyse the comparison details below:
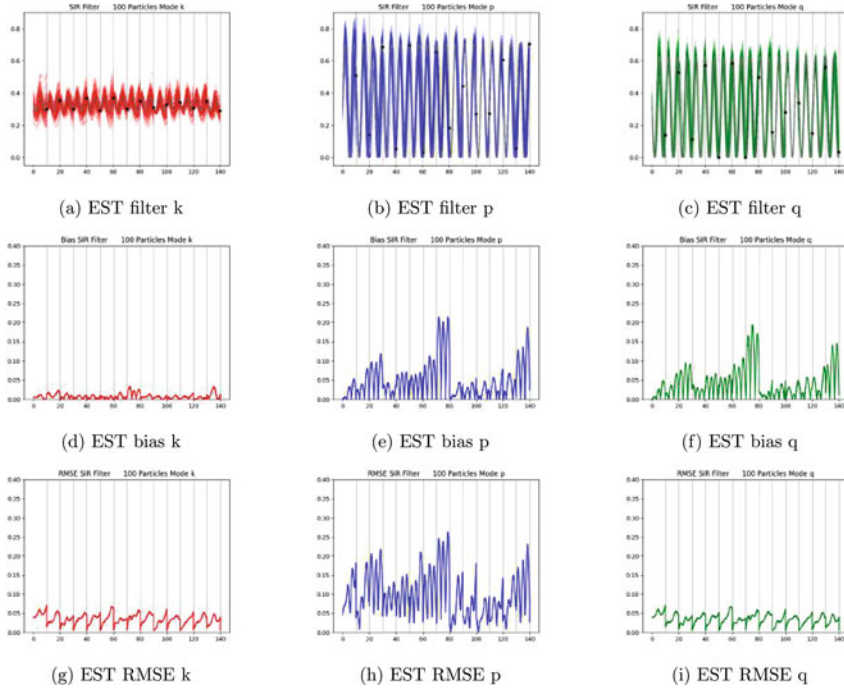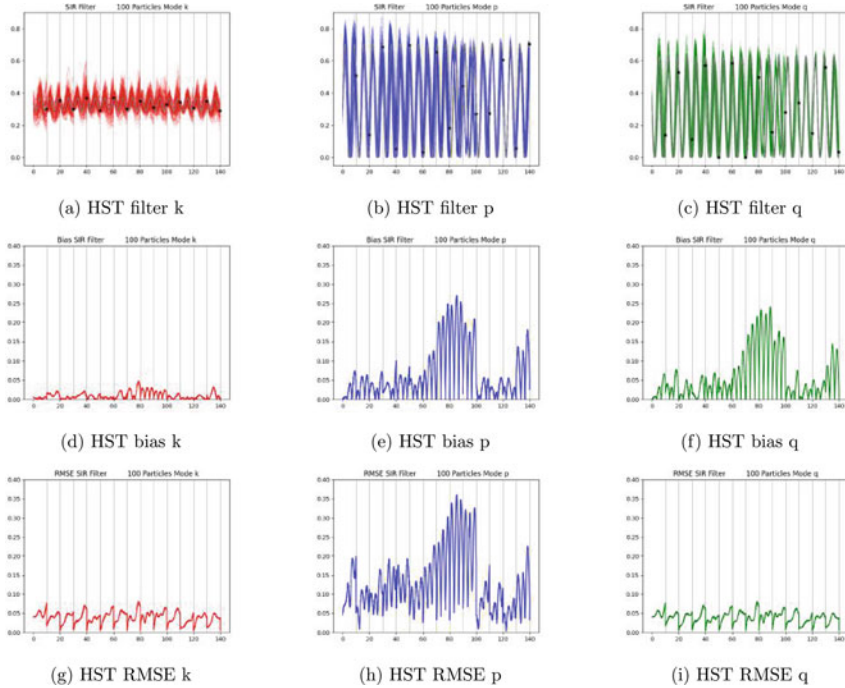
Fig. 6 Filtering experiment for EST model using SIR particle filter. (**a–c**) Ensemble evolution for 100 particles in mode k (**a**, red), p (**b**, blue), and q (**c**, green). The signal (grey) is the deterministic model and the observations (black stars) are noisy and taken and assimilated every 10 time units. (**d–f**) Bias of the filtering ensemble. (**g–h**) RMSE of the filtering ensemble

Mode k

This is the least energetic of all the modes (hence the reason why we observe it with the least amount of measurement noise). The cloud of particles is well placed around the truth even with the small sample. The bias remains small for both the HST and the EST versions and it reduces significantly when observations are assimilated more frequently (see Fig. 9 in Appendix 3) as well as when we use a large number (500) of particles (see Fig. 10 in Appendix 3). The RMSE remains small in all cases and decreases (though not substantially) when the DA step is small.

Modes p and q

These are the two energetic modes of the system. We used here a measurement noise that is one order of magnitude larger. Despite this, the results remain equally good. The cloud of particles provide a good envelope for the truth at all times. This validates the choice of the stochasticity: the uncertainty is properly modelled.

Fig. 7 Filtering experiment for HST model using SIR particle filter. (**a–c**) Ensemble evolution for 100 particles in mode k (**a**, red), p (**b**, blue), and q (**c**, green). The signal (grey) is the deterministic model and the observations (black stars) are noisy and taken and assimilated every 10 time units. (**d–f**) Bias of the filtering ensemble. (**g–h**) RMSE of the filtering ensemble

For both models the bias can become very large, reaching 30% of the size of the oscillations for the HST model and 25% of the size of the oscillations for the EST model. As expected, it is drastically reduced when observations are assimilated more frequently. The RMSE for mode p is also large but substantially smaller for mode q. The addition of more intermediate DA steps or more particles has a less pronounced effect for the q mode.

Remark 2 We record the Effective Sample Size (ESS) for a typical run (for both EST and HST) in Fig. 11. As usual, the ESS is computed just before the application of the resampling procedure. The ESS is seen to decay dramatically from 100 down to single digits numbers in most instances in time.

Remark 3 We record the results over 10 independent runs of the filtering experiment for the EST model with 500 ensemble members in Fig. 12. More precisely, in graphs 12a, b and c each, we plot the mean across the 10 independent filtering runs together with the evolution of the signal and the individual ensemble means for each mode. The mean bias as well as the envelope obtained from the independent runs are shown in graphs 12d, e and f. The same is shown for the RMSE in graphs 12g,

h and i. Compared with a single run of the same experiment reported in Fig. 10 we observe the approximations are now near perfect (the statistical error has been drastically reduced).

4 Conclusions

The introduction of stochasticity into the deterministic triad models leads to two new stochastic models. Stochasticity is introduced in a principled way (rather than ad-hoc). It starts with a full scale fluid dynamic model which is randomly perturbed. At the full scale, the stochastic parametrisation models the small-scale effects in fluid dynamics modelling. In particular, it efficiently captures the high-frequency small-scale dynamics and correctly correlates it with the slow, large-scale fluid motion. In addition, it is constrained to conserve either the helicity or the kinetic energy of the system. This inspires two different stochastic triad models of Euler type which we compare using data assimilation procedures based on particle filtering. The methodology we employ can be used as a benchmark when analysing new types of stochastic parametrisations: ours is the first study that assesses the efficiency of stochastic parametrisations from a data assimilation perspective.

The introduction of stochasticity ensures that the correct spread (one that preserves the physical properties of the system) is introduced in the ensemble of particles. In its absence, the particle filter degenerates quite rapidly: after a few DA steps, a single particle survives the culling procedure which does offer a good approximation of the truth. A purely deterministic transition kernel does not work, generating a rapid degeneracy of the particle filter.

The two stochastic systems (one which preserves helicity and the other one which preserves energy) are analysed using a standard particle filter. There is no need for additional procedures (such as tempering, nudging, or jittering). They perform equally well from the viewpoint of DA: both the RMSE and the bias are drastically reduced and stabilised when the noise is carefully calibrated. The two different stochastic kernels require different noise calibrations in order to perform well in similar data assimilation scenarios. This is somehow expected, given that the underlying stochastic parametrisations preserve different physical quantities.

Appendix 1: Notation and Basic Identities

Notation

In this work we use the following notation. We write $\mathbb{Z}$, $\mathbb{R}$ and $\mathbb{C}$ for the sets of integers, real numbers and complex numbers, respectively. Boldface letters denote three-dimensional complex vectors. For two complex vectors **a** and **b** with

components a_j and b_j, their dot product is denoted by

$$\mathbf{a} \cdot \mathbf{b} = \sum_{j=1}^{3} a_j b_j = a_j b_j \in \mathbb{C}.$$

This paper uses the Einstein convention of summing over repeated indices. The norm of the complex vector $\mathbf{a}$ is defined as

$$|\mathbf{a}| = \sqrt{\mathbf{a} \cdot \mathbf{a}^*} = \sqrt{a_j a_j^*} \geq 0,$$

with the superscript symbol * denoting complex conjugation. Further, the cross product of two vectors $\mathbf{a}$ and $\mathbf{b}$ is given by

$$\mathbf{a} \times \mathbf{b} = (a_2 b_3 - a_3 b_2, a_3 b_1 - a_1 b_3, a_1 b_2 - a_2 b_1) \in \mathbb{C}^3.$$

The gradient of a scalar field $\phi : \mathcal{D} \subseteq \mathbb{R}^3 \to \mathbb{C}$ at a point $\mathbf{x} \in \mathcal{D}$ is denoted by

$$\nabla \phi(\mathbf{x}) = \big(\partial_1 \phi(\mathbf{x}), \partial_2 \phi(\mathbf{x}), \partial_3 \phi(\mathbf{x})\big) \in \mathbb{C}^3.$$

The divergence of a vector field $\boldsymbol{\psi} : \mathcal{D} \subseteq \mathbb{R}^3 \to \mathbb{C}^3$ with components ψ_j at a point $\mathbf{x} \in \mathcal{D}$ is defined as

$$\nabla \cdot \boldsymbol{\psi}(\mathbf{x}) = \partial_1 \psi_1(\mathbf{x}) + \partial_2 \psi_2(\mathbf{x}) + \partial_3 \psi_3(\mathbf{x}) = \partial_j \psi_j(\mathbf{x}) \in \mathbb{C}$$

and the curl of $\boldsymbol{\psi}$ at $\mathbf{x}$ is given by

$$\nabla \times \boldsymbol{\psi}(\mathbf{x}) = \big(\partial_2 \psi_3(\mathbf{x}) - \partial_3 \psi_2(\mathbf{x}), \partial_3 \psi_1(\mathbf{x}) - \partial_1 \psi_3(\mathbf{x}), \partial_1 \psi_2(\mathbf{x}) - \partial_2 \psi_1(\mathbf{x})\big) \in \mathbb{C}^3.$$

Vector Identities

For three vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$, we have the following algebraic vector identities

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}),$$

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c},$$

$$(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) = (\mathbf{a} \cdot \mathbf{c})(\mathbf{b} \cdot \mathbf{d}) - (\mathbf{b} \cdot \mathbf{c})(\mathbf{a} \cdot \mathbf{d}),$$

$$\mathbf{a} \times \mathbf{a} = \mathbf{0}.$$

Moreover, we have the vector calculus identity

$$(\mathbf{a} \cdot \nabla)\mathbf{b} + b_j \nabla a_j = -\mathbf{a} \times (\nabla \times \mathbf{b}) + \nabla(\mathbf{a} \cdot \mathbf{b}). \tag{56}$$

Appendix 2: Derivation of Triad Models

Deterministic Euler

We compute the projection of the terms corresponding to the deterministic Euler
vorticity equation in curl form:

$$\partial_t \mathbf{v} - \mathbf{v} \times (\nabla \times \mathbf{v}) + \tfrac{1}{2}\nabla|\mathbf{v}|^2 = -\nabla p. \tag{57}$$

onto the helical basis. For the time-derivative we get

$$\int_{\mathcal{D}} \partial_t \mathbf{v}(\mathbf{x}, t) \cdot \mathbf{h}_{s_k}^*(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{x}}\, d\mathbf{x} = \sum_{\mathbf{p}} \sum_{s_p} \partial_t a_{s_p}(\mathbf{p}, t) \mathbf{h}_{s_p}(\mathbf{p}) \cdot \mathbf{h}_{s_k}^*(\mathbf{k}) \int_{\mathcal{D}} e^{i(\mathbf{p}-\mathbf{k})\cdot\mathbf{x}}\, d\mathbf{x}$$

$$= L^3 \sum_{s_p} \partial_t a_{s_p}(\mathbf{k}, t) \mathbf{h}_{s_p}(\mathbf{k}) \cdot \mathbf{h}_{s_k}^*(\mathbf{k})$$

$$= L^3 \partial_t a_{s_k}(\mathbf{k}, t) \mathbf{h}_{s_k}(\mathbf{k}) \cdot \mathbf{h}_{s_k}^*(\mathbf{k})$$

$$= 2L^3 \partial_t a_{s_k}(\mathbf{k}, t).$$

The vorticity term gives

$$\int_{\mathcal{D}} (\mathbf{v}(\mathbf{x}, t) \times \boldsymbol{\omega}(\mathbf{x}, t)) \cdot \mathbf{h}_{s_k}^*(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{x}}\, d\mathbf{x}$$

$$= \sum_{\mathbf{p},\mathbf{q}} \sum_{s_p, s_q} a_{s_p}(\mathbf{p}, t) s_q |\mathbf{q}| a_{s_q}(\mathbf{q}, t) \mathbf{h}_{s_p}(\mathbf{p}) \times \mathbf{h}_{s_q}(\mathbf{q}) \cdot \mathbf{h}_{s_k}^*(\mathbf{k}) \int_{\mathcal{D}} e^{i(\mathbf{p}+\mathbf{q}-\mathbf{k})\cdot\mathbf{x}}\, d\mathbf{x}$$

$$= L^3 \sum_{\mathbf{p}+\mathbf{q}+\mathbf{k}=0} \sum_{s_p, s_q} a_{s_p}^*(\mathbf{p}, t) s_q |\mathbf{q}| a_{s_q}^*(\mathbf{q}, t) \mathbf{h}_{s_p}^*(\mathbf{p}) \times \mathbf{h}_{s_q}^*(\mathbf{q}) \cdot \mathbf{h}_{s_k}^*(\mathbf{k}). \tag{58}$$

Note that we can write (58) in a form which is symmetric in **p** and **q** since, renaming
the indices,

$$\sum_{\mathbf{p}+\mathbf{q}+\mathbf{k}=0} \sum_{s_p, s_q} a_{s_p}^*(\mathbf{p}, t) s_q |\mathbf{q}| a_{s_q}^*(\mathbf{q}, t) \mathbf{h}_{s_p}^*(\mathbf{p}) \times \mathbf{h}_{s_q}^*(\mathbf{q}) \cdot \mathbf{h}_{s_k}^*(\mathbf{k})$$

$$= \sum_{\mathbf{p}+\mathbf{q}+\mathbf{k}=0} \sum_{s_p, s_q} a_{s_q}^*(\mathbf{q}, t) s_p |\mathbf{p}| a_{s_p}^*(\mathbf{p}, t) \mathbf{h}_{s_q}^*(\mathbf{q}) \times \mathbf{h}_{s_p}^*(\mathbf{p}) \cdot \mathbf{h}_{s_k}^*(\mathbf{k})$$

$$= -\sum_{\mathbf{p}+\mathbf{q}+\mathbf{k}=0} \sum_{s_p, s_q} a_{s_q}^*(\mathbf{q}, t) s_p |\mathbf{p}| a_{s_p}^*(\mathbf{p}, t) \mathbf{h}_{s_p}^*(\mathbf{p}) \times \mathbf{h}_{s_q}^*(\mathbf{q}) \cdot \mathbf{h}_{s_k}^*(\mathbf{k}).$$

Therefore,

$$\int_{\mathcal{D}} (\mathbf{v}(\mathbf{x}, t) \times \boldsymbol{\omega}(\mathbf{x}, t)) \cdot \mathbf{h}_{s_k}^*(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{x}} \, d\mathbf{x} \tag{59}$$

$$= \frac{L^3}{2} \sum_{\mathbf{p}+\mathbf{q}+\mathbf{k}=0} \sum_{s_p, s_q} (s_q|\mathbf{q}| - s_p|\mathbf{p}|) a_{s_p}^*(\mathbf{p}, t) a_{s_q}^*(\mathbf{q}, t) \mathbf{h}_{s_p}^*(\mathbf{p}) \times \mathbf{h}_{s_q}^*(\mathbf{q}) \cdot \mathbf{h}_{s_k}^*(\mathbf{k}).$$

$$\tag{60}$$

Moreover, the gradient terms in (57) vanish upon expansion into helical modes. Thus, the Euler equation (57) in helical basis becomes

$$\partial_t a_{s_k}(\mathbf{k}, t) = -\frac{1}{4} \sum_{\mathbf{p}+\mathbf{q}+\mathbf{k}=0} \sum_{s_p, s_q} (s_p|\mathbf{p}| - s_q|\mathbf{q}|) a_{s_p}^*(\mathbf{p}, t) a_{s_q}^*(\mathbf{q}, t) \mathbf{h}_{s_p}^*(\mathbf{p}) \times \mathbf{h}_{s_q}^*(\mathbf{q}) \cdot \mathbf{h}_{s_k}^*(\mathbf{k}).$$

SALT Euler

We expand the 3D SALT Navier-Stokes equation (23) using (6) and (37). Assume $b_s(-\mathbf{p}) = b_s^*(\mathbf{p})$.

$$\int_{\mathcal{D}} d\mathbf{v}(\mathbf{x}, t) \cdot \mathbf{h}_{s_k}^*(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{x}} \, d\mathbf{x} = \sum_{\mathbf{p}} \sum_{s_p} da_{s_p}(\mathbf{p}, t) \mathbf{h}_{s_p}(\mathbf{p}) \cdot \mathbf{h}_{s_k}^*(\mathbf{k}) \int_{\mathcal{D}} e^{i(\mathbf{p}-\mathbf{k})\cdot\mathbf{x}} \, d\mathbf{x}$$

$$= 2L^3 da_{s_k}(\mathbf{k}, t)$$

And

$$\int_{\mathcal{D}} d\mathbf{x}_t(\mathbf{x}, t) \times \boldsymbol{\omega}(\mathbf{x}, t) \cdot \mathbf{h}_{s_k}^*(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{x}} \, d\mathbf{x}$$

$$= \sum_{\mathbf{p}, \mathbf{q}} \sum_{s_p, s_q} \left[a_{s_p}(\mathbf{p}, t) dt + b_{s_p}(\mathbf{p}) \circ dW_t \right] s_q|\mathbf{q}| a_{s_q}(\mathbf{q}, t) \mathbf{h}_{s_p}(\mathbf{p})$$

$$\times \mathbf{h}_{s_q}(\mathbf{q}) \cdot \mathbf{h}_{s_k}^*(\mathbf{k}) \int_{\mathcal{D}} e^{i(\mathbf{p}+\mathbf{q}-\mathbf{k})\cdot\mathbf{x}} \, d\mathbf{x}$$

$$= L^3 \sum_{\mathbf{p}+\mathbf{q}+\mathbf{k}=0} \sum_{s_p, s_q} \left[a_{s_p}^*(\mathbf{p}, t) dt + b_{s_p}^*(\mathbf{p}) \circ dW_t \right] s_q|\mathbf{q}| a_{s_q}^*(\mathbf{q}, t) \mathbf{h}_{s_p}^*(\mathbf{p})$$

$$\times \mathbf{h}_{s_q}^*(\mathbf{q}) \cdot \mathbf{h}_{s_k}^*(\mathbf{k}).$$

Renaming the indices, we can write

$$L^3 \sum_{\mathbf{p}+\mathbf{q}+\mathbf{k}=0} \sum_{s_p,s_q} \left[a_{s_p}^*(\mathbf{p},t)\mathrm{d}t + b_{s_p}^*(\mathbf{p}) \circ \mathrm{d}W_t\right] s_q |\mathbf{q}| a_{s_q}^*(\mathbf{q},t) \mathbf{h}_{s_p}^*(\mathbf{p}) \times \mathbf{h}_{s_q}^*(\mathbf{q}) \cdot \mathbf{h}_{s_k}^*(\mathbf{k})$$

$$=L^3 \sum_{\mathbf{p}+\mathbf{q}+\mathbf{k}=0} \sum_{s_p,s_q} \left[a_{s_q}^*(\mathbf{q},t)\mathrm{d}t + b_{s_q}^*(\mathbf{q}) \circ \mathrm{d}W_t\right] s_p |\mathbf{p}| a_{s_p}^*(\mathbf{p},t) \mathbf{h}_{s_q}^*(\mathbf{q}) \times \mathbf{h}_{s_p}^*(\mathbf{p}) \cdot \mathbf{h}_{s_k}^*(\mathbf{k})$$

$$=-L^3 \sum_{\mathbf{p}+\mathbf{q}+\mathbf{k}=0} \sum_{s_p,s_q} \left[a_{s_q}^*(\mathbf{q},t)\mathrm{d}t + b_{s_q}^*(\mathbf{q}) \circ \mathrm{d}W_t\right] s_p |\mathbf{p}| a_{s_p}^*(\mathbf{p},t) \mathbf{h}_{s_p}^*(\mathbf{p}) \times \mathbf{h}_{s_q}^*(\mathbf{q}) \cdot \mathbf{h}_{s_k}^*(\mathbf{k}).$$

Thus, we arrive at

$$\int_{\mathcal{D}} \mathrm{d}\mathbf{x}_t(\mathbf{x},t) \times \boldsymbol{\omega}(\mathbf{x},t) \cdot \mathbf{h}_{s_k}^*(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{x}} \, \mathrm{d}\mathbf{x}$$

$$= \frac{L^3}{2} \sum_{\mathbf{p}+\mathbf{q}+\mathbf{k}=0} \sum_{s_p,s_q} \left[(s_q |\mathbf{q}| b_{s_p}^*(\mathbf{p}) \circ \mathrm{d}W_t a_{s_q}^*(\mathbf{q},t) - s_p |\mathbf{p}| b_{s_q}^*(\mathbf{q}) \circ \mathrm{d}W_t a_{s_p}^*(\mathbf{p},t)) \right.$$

$$\left. + (a_{s_p}^*(\mathbf{p},t)\mathrm{d}t s_q |\mathbf{q}| a_{s_q}^*(\mathbf{q},t) - a_{s_q}^*(\mathbf{q},t)\mathrm{d}t s_p |\mathbf{p}| a_{s_p}^*(\mathbf{p},t)) \right] \mathbf{h}_{s_p}^*(\mathbf{p}) \times \mathbf{h}_{s_q}^*(\mathbf{q}) \cdot \mathbf{h}_{s_k}^*(\mathbf{k}).$$

Therefore,

$$\mathrm{d}a_{s_k}(\mathbf{k},t) = \frac{1}{4} \sum_{\mathbf{p}+\mathbf{q}+\mathbf{k}=0} \sum_{s_p,s_q} \left[(s_q |\mathbf{q}| b_{s_p}^*(\mathbf{p}) \circ \mathrm{d}W_t a_{s_q}^*(\mathbf{q},t) - s_p |\mathbf{p}| b_{s_q}^*(\mathbf{q}) \circ \mathrm{d}W_t a_{s_p}^*(\mathbf{p},t)) \right.$$

$$\left. + (a_{s_p}^*(\mathbf{p},t)\mathrm{d}t s_q |\mathbf{q}| a_{s_q}^*(\mathbf{q},t) - a_{s_q}^*(\mathbf{q},t)\mathrm{d}t s_p |\mathbf{p}| a_{s_p}^*(\mathbf{p},t)) \right] \mathbf{h}_{s_p}^*(\mathbf{p}) \times \mathbf{h}_{s_q}^*(\mathbf{q}) \cdot \mathbf{h}_{s_k}^*(\mathbf{k}).$$

LU Euler

Written in terms of the SALT model, the LU model is

$$\mathrm{d}\mathbf{v} + \mathrm{d}\mathbf{x}_t \cdot \nabla \mathbf{v} + \mathbf{v}^j \nabla \mathrm{d}\mathbf{x}_t^j - \mathbf{v}^j \nabla(\boldsymbol{\xi} \circ \mathrm{d}W_t)^j = -\nabla \, \mathrm{d}p.$$

Expanding the additional term gives

$$\nabla(\boldsymbol{\xi} \circ \mathrm{d}W_t)^j = \nabla \sum_{\mathbf{q}} (b_\pm(\mathbf{q})\mathbf{h}_\pm(\mathbf{q})e^{i\mathbf{q}\cdot\mathbf{x}} \circ \mathrm{d}W_t)^j = \sum_{\mathbf{q}} i\mathbf{q} b_\pm(\mathbf{q})\mathbf{h}_\pm^j(\mathbf{q})e^{i\mathbf{q}\cdot\mathbf{x}} \circ \mathrm{d}W_t.$$

Thus we get

$$\mathbf{v}^j \nabla(\boldsymbol{\xi} \circ dW_t)^j = \sum_{\mathbf{p},\mathbf{q}} \sum_{s_p,s_q} a_{s_p}(\mathbf{p},t) \mathbf{h}_{s_p}^j(\mathbf{p}) i\mathbf{q} b_{s_q}(\mathbf{q}) \mathbf{h}_{s_q}^j(\mathbf{q}) e^{i(\mathbf{p}+\mathbf{q})\cdot\mathbf{x}} \circ dW_t.$$

Now projecting and renaming, we have

$$\int_{\mathcal{D}} \mathbf{v}^j \nabla(\boldsymbol{\xi} \circ dW_t)^j \cdot a_{s_k}^*(\mathbf{k},t) \mathbf{h}_{s_k}^*(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{x}} \, d\mathbf{x} =$$

$$= \sum_{\mathbf{p},\mathbf{q}} \sum_{s_p,s_q} a_{s_k}^*(\mathbf{k},t) a_{s_p}(\mathbf{p},t) [b_{s_q}(\mathbf{q}) \circ dW_t] (i\mathbf{q} \cdot \mathbf{h}_{s_k}^*(\mathbf{k})) (\mathbf{h}_{s_p}(\mathbf{p}) \cdot \mathbf{h}_{s_q}(\mathbf{q})) \int_{\mathcal{D}} e^{(\mathbf{p}+\mathbf{q}-\mathbf{k})\cdot\mathbf{x}} \, d\mathbf{x}$$

$$= L^3 \sum_{\mathbf{p}+\mathbf{q}+\mathbf{k}=0} \sum_{s_p,s_q} a_{s_k}^*(\mathbf{k},t) a_{s_p}^*(\mathbf{p},t) [b_{s_q}^*(\mathbf{q}) \circ dW_t] (-i\mathbf{q} \cdot \mathbf{h}_{s_k}^*(\mathbf{k})) (\mathbf{h}_{s_p}^*(\mathbf{p}) \cdot \mathbf{h}_{s_q}^*(\mathbf{q}))$$

$$= L^3 \sum_{\mathbf{p}+\mathbf{q}+\mathbf{k}=0} \sum_{s_p,s_q} (f_k^{pq})^* a_{s_k}^*(\mathbf{k},t) (a_{s_p}^*(\mathbf{p},t) [b_{s_q}^*(\mathbf{q}) \circ dW_t] + a_{s_q}^*(\mathbf{q},t) [b_{s_p}^*(\mathbf{p}) \circ dW_t])$$

with

$$f_k^{pq} = (-i(\mathbf{p}+\mathbf{q}) \cdot \mathbf{h}_{s_k}(\mathbf{k})) (\mathbf{h}_{s_p}(\mathbf{p}) \cdot \mathbf{h}_{s_q}(\mathbf{q})).$$

Note that, due to the triad condition $\mathbf{p} + \mathbf{q} + \mathbf{k} = 0$,

$$f_k^{pq} = f_p^{kq} = f_q^{kp} = 0$$

so that the difference term between SALT and LU vanishes in the helical projection and the two projected models coincide.

Appendix 3: Supplementary Numerics

Calibration of the Noise Amplitude

To calibrate the noise amplitude for the data assimilation experiments, we rely on two forecast verification metrics. The rank histogram (or Talagrand histogram) and the continuous ranked probability score (CRPS). We evaluated these for both models on 64 different noise amplitude vectors. The metrics were recorded by running the data assimilation/particle filtering experiment described in the main text for the noise amplitude vectors $\mathbf{b} = [b_k, b_p, b_q]$ resulting from all possible combinations of $b_k \in \{0.05, 0.1, 0.2, 0.5\}$, $b_p \in \{0.025, 0.05, 0.1, 0.2\}$ and $b_q \in \{0.01, 0.02, 0.04, 0.1\}$. We present the mean CRPS scores for top 5 tested noise

Table 1 CRPS scores (lower is better) for the five best—in terms of overall mean CRPS score—tested noise amplitude vectors **b** (first column). The mean CRPS score over the three modal energies is shown for the EST (second column) and HST (third column) models. The overall mean CRPS score for the respective noise amplitude is shown in the *Mean* column (fourth column). Finally, we provide references to the figures showing the associated rank histograms (last column)

b	EST	HST	Mean	Histograms
[0.05, 0.025, 0.01]	0.0282	0.0394	0.0338	Figure 8a + f
[0.05, 0.025, 0.02]	0.0338	0.0392	0.0365	Figure 8b + g
[0.10, 0.025, 0.02]	0.0352	0.0385	0.0368	Figure 8c + h
[0.10, 0.05, 0.02]	0.0393	0.0356	0.0375	Figure 8d + i
[0.10, 0.05, 0.01]	0.0363	0.0409	0.0386	Figure 8e + j

(a) EST   (b) EST   (c) EST   (d) EST   (e) EST

(f) HST   (g) HST   (h) HST   (i) HST   (j) HST
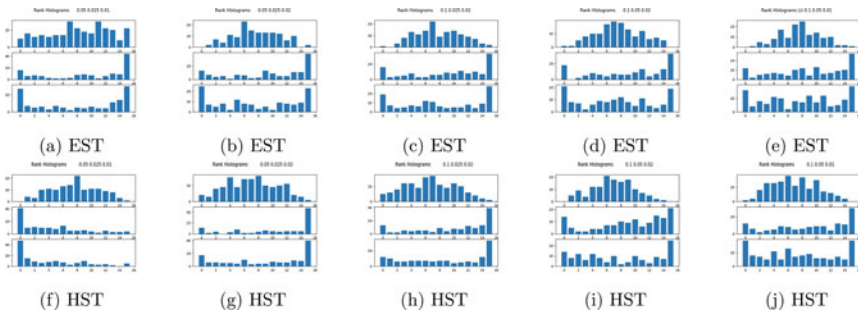
Fig. 8 Rank histograms for the 5 best noise amplitude vectors in terms of CRPS score (see Table 1) for the EST and HST models. Each individual graph shows the rank histogram of an ensemble of 15 particles run up to a final time of 1400, with data assimilation performed every 10 time units. The top histogram in each subfigure represents the ensemble for mode **k**, the middle histogram represents the ensemble for mode **p** and the bottom one for mode **q**

amplitude vectors in Table 1. The mean is taken across both models. Based on this calibration, we chose the case **b** = [0.10, 0.05, 0.01] for the data assimilation experiment, as the other ones have inferior rank histograms, so we achieve a good balance between the visual judgment of rank histograms and the CRPS score (Fig. 8).

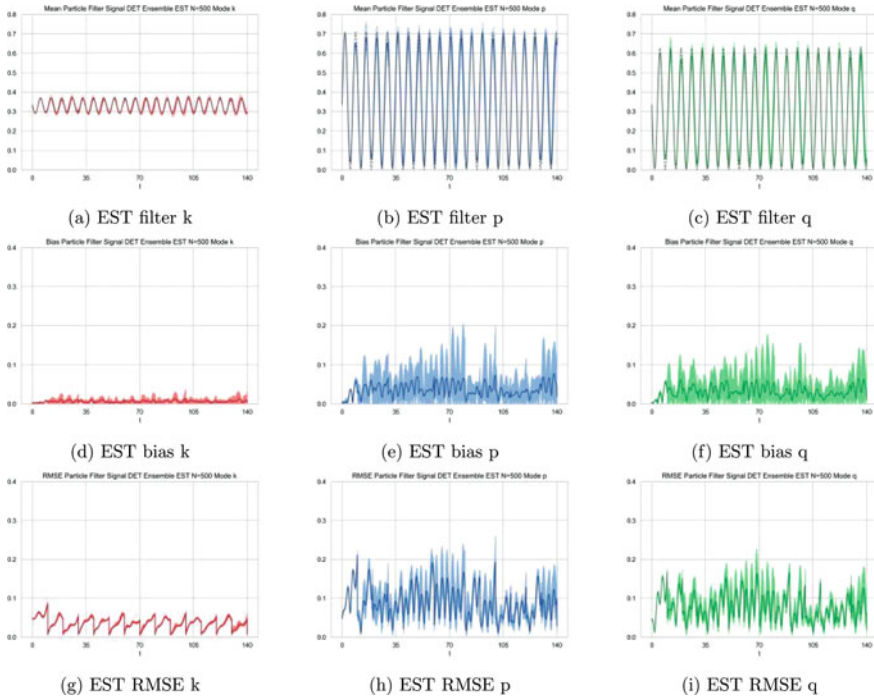## *Data Assimilation Verification*

See Figs. 9, 10, 11, and 12.
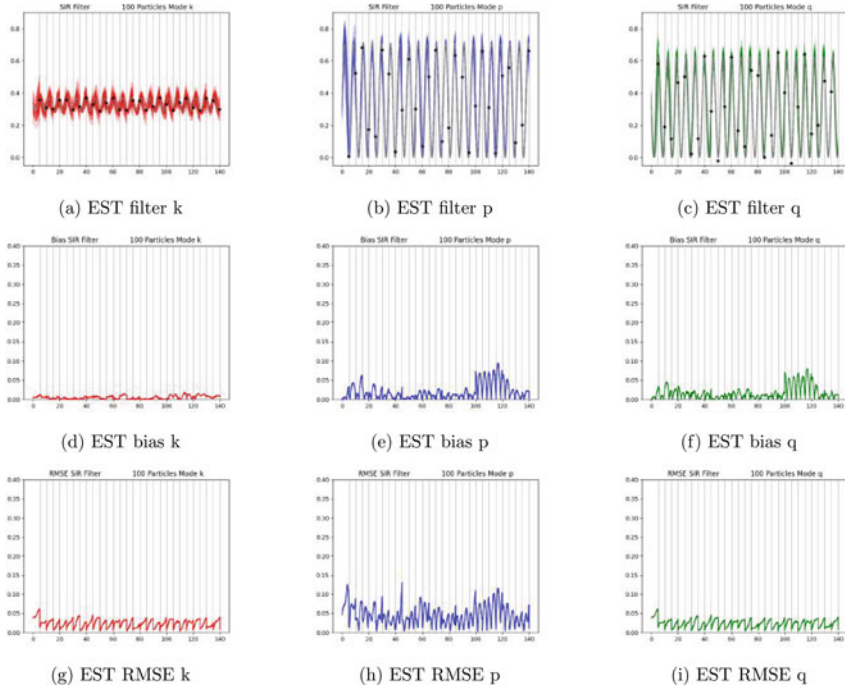
Fig. 9 Filtering experiment for EST model using SIR particle filter with a small data assimilation interval of 5 time units. (**a–c**) Ensemble evolution for 100 particles in mode k (**a**, red), p (**b**, blue), and q (**c**, green). The signal (grey) is the deterministic model and the observations (black stars) are noisy and taken and assimilated every 5 time units. (**d–f**) Bias of the filtering ensemble. (**g–h**) RMSE of the filtering ensemble
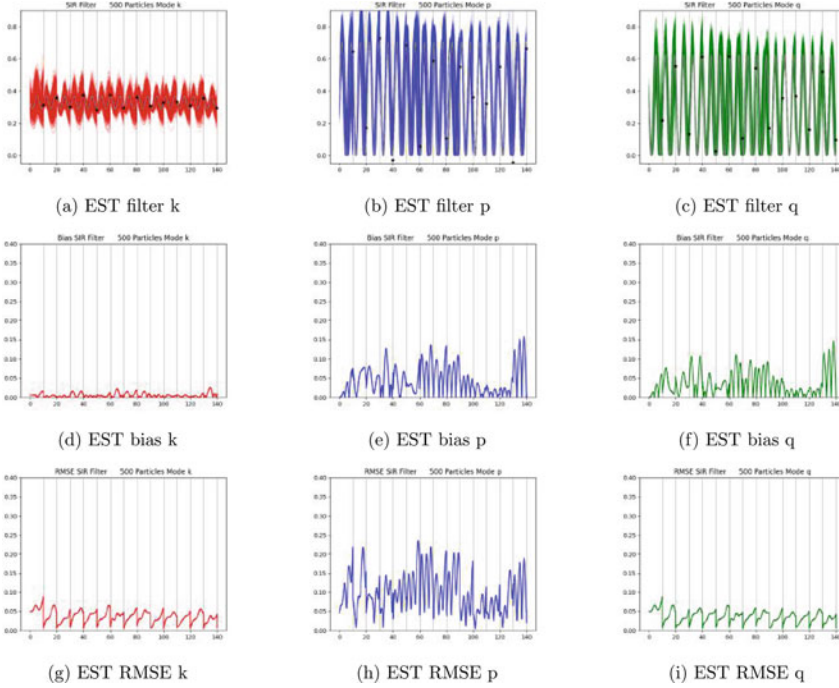
Fig. 10 Filtering experiment for EST model using SIR particle filter with a large particle ensemble of 500 members. (**a–c**) Ensemble evolution for 500 particles in mode k (**a**, red), p (**b**, blue), and q (**c**, green). The signal (grey) is the deterministic model and the observations (black stars) are noisy and taken and assimilated every 5 time units. (**d–f**) Bias of the filtering ensemble. (**g–h**) RMSE of the filtering ensemble
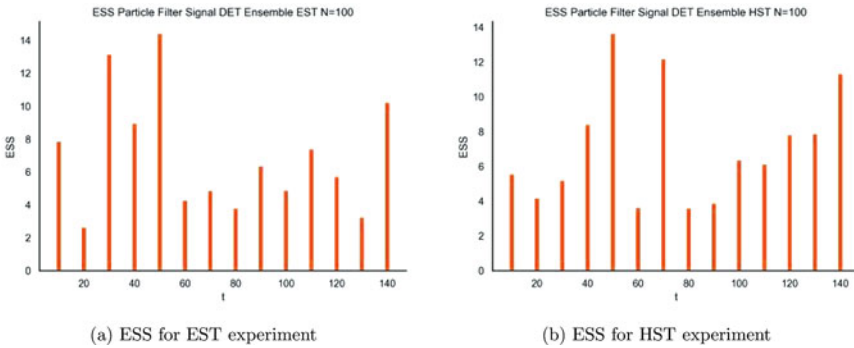
Fig. 11 Typical ESS for the filtering experiments. See Remark 2 in the main text. (**a**) ESS for EST experiment. (**b**) ESS for HST experiment

(a) EST filter k   (b) EST filter p   (c) EST filter q

(d) EST bias k   (e) EST bias p   (f) EST bias q

(g) EST RMSE k   (h) EST RMSE p   (i) EST RMSE q

Fig. 12 Statistics of the mean ensemble over 10 independent runs of the particle filter with 500 ensemble members for the EST model. See Remark 3 in the main text

References

[1] G. Tadmor B. Noack, M. Morzinsky. *Reduced-Order Modelling for Flow Control*, volume 528. CISM Courses and Lectures Springer, 2010.

[2] Alan Bain and Dan Crisan. *Fundamentals of stochastic filtering*, volume 60 of *Stochastic Modelling and Applied Probability*. Springer, New York, 2009. Available from: https://doi.org/10.1007/978-0-387-76896-0.

[3] Carl M Bender, Darryl D Holm, and Daniel W Hook. Complexified dynamical systems. *Journal of Physics A: Mathematical and Theoretical*, 40(32):F793, 2007.

[4] Judith Berner, Ulrich Achatz, Lauriane Batte, Lisa Bengtsson, Alvaro De La Camara, Hannah M Christensen, Matteo Colangeli, Danielle RB Coleman, Daan Crommelin, Stamen I Dolaptchiev, et al. Stochastic parameterization: Toward a new view of weather and climate models. *Bulletin of the American Meteorological Society*, 98(3):565–588, 2017.

[5] L. Biferale. Shell models of energy cascade in turbulence. *Annual Review of Fluid Mechanics*, 35:441–468, 2003.

[6] L. Biferale. Shell models of energy cascade in turbulence. *Annual Rev. Fluid Mech.*, 35:441–468, 2003.

[7] Tomas Bohr, Mogens H Jensen, Giovanni Paladin, and Angelo Vulpiani. *Dynamical systems approach to turbulence*. CUP, 1998.

[8] B. Chapron, P. Dérian, E. Mémin, and V. Resseguier. Large-scale flows under location uncertainty: a consistent stochastic framework. *Quarterly Journal of the Royal Meteorological Society*, 144(710):251–260, 2018. Available from: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3198, arXiv:https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.3198, https://doi.org/10.1002/qj.3198.

[9] Qiaoning Chen, Shiyi Chen, and Gregory L Eyink. The joint cascade of energy and helicity in three-dimensional turbulence. *Physics of Fluids*, 15(2):361–374, 2003. Available from: https://doi.org/10.1063/1.1533070.

[10] Qiaoning Chen, Shiyi Chen, Gregory L Eyink, and Darryl D Holm. Intermittency in the joint cascade of energy and helicity. *Physical Review Letters*, 90(21):214503, 2003. Available from: https://doi.org/10.1103/PhysRevLett.90.214503.

[11] Colin J. Cotter, Dan Crisan, Darryl D. Holm, Wei Pan, and Igor Shevchenko. Numerically modelling stochastic lie transport in fluid dynamics, 2018. Available from: https://arxiv.org/abs/1801.09729, https://doi.org/10.48550/ARXIV.1801.09729.

[12] Alex DD Craik. *Wave interactions and fluid flows*. Cambridge Nonlinear Sciences, 1998.

[13] B. Geurts, D. Holm, and E. Luesink. Lyapunov exponents of two stochastic lorenz 63 systems. *Journal of statistical physics*, 179:1343–1365, 2017.

[14] Bernard J. Geurts, Darryl D. Holm, and Erwin Luesink. Lyapunov exponents of two stochastic Lorenz 63 systems. *Journal of Statistical Physics*, 179(5):1343–1365, Jun 2020. Available from: https://doi.org/10.1007/s10955-019-02457-3.

[15] E. B. Gledzer. Rotation and helicity effects in cascade models of turbulence. *Doklady Physics*, 53:216–220, 2008.

[16] Darryl D. Holm. Variational principles for stochastic fluid dynamics. *Proc. A.*, 471(2176):20140963, 19, 2015.

[17] Darryl D Holm and Ruiao Hu. Stochastic effects of waves on currents in the ocean mixed layer. *Journal of Mathematical Physics*, 62(7):073102, 2021.

[18] Darryl D Holm, Jerrold E Marsden, and Tudor S Ratiu. The Euler–Poincaré equations and semidirect products with applications to continuum theories. *Advances in Mathematics*, 137(1):1–81, 1998.

[19] Ioannis Karatzas and Steven Shreve. *Brownian motion and stochastic calculus*, volume 113. Springer Science & Business Media, 2012.

[20] Robert H Kraichnan. Small-scale structure of a scalar field convected by turbulence. *The Physics of Fluids*, 11(5):945–953, 1968.

[21] Long Li. *Stochastic modeling and numerical simulation of ocean dynamics*. PhD thesis, Université de Rennes 1, 2021.

[22] Alexander Lobbe. alobbe/stochastic-triads: Initial release, April 2023. Available from: https://doi.org/10.5281/zenodo.7845270.

[23] Etienne Mémin. Fluid flow dynamics under location uncertainty. *Geophys. Astrophys. Fluid Dyn.*, 108(2):119–146, 2014.

[24] E. A. Novikov. *Infinitely divisible distributions in turbulence*. Physical Review E, 1994.

[25] Valentin Resseguier, Agustin M. Picard, Etienne Memin, and Bertrand Chapron. Quantifying truncation-related uncertainties in unsteady fluid dynamics reduced order models. *SIAM/ASA Journal on Uncertainty Quantification*, 9(3):1152–1183, 2021. Available from: https://doi.org/10.1137/19M1354819.

[26] Peter Jan van Leeuwen, Dan Crisan, Oana Lang, and Roland Potthast. Bayesian inference for fluid dynamics: A case study for the stochastic rotating shallow water model, 2021. arXiv:2112.15216.

[27] Dario Vincenzi and John D Gibbon. How close are shell models to the 3d Navier–Stokes equations? *Nonlinearity*, 34(8):5821, 2021.

[28] Fabian Waleffe. The nature of triad interactions in homogeneous turbulence. *Physics of Fluids A: Fluid Dynamics*, 4(2):350–363, 1992.

[29] Fabian Waleffe. Inertial transfers in the helical decomposition. *Physics of Fluids A: Fluid Dynamics*, 5(3):677–685, 1993. Available from: https://doi.org/10.1063/1.858651.