Appl. Math. Mech. -Engl. Ed., **44**(7), 1175–1198 (2023)

APPLIED MATHEMATICS AND MECHANICS (ENGLISH EDITION)

https://doi.org/10.1007/s10483-023-2990-9

Gaussian process hydrodynamics*

H. OWHADI†

California Institute of Technology, MC 9-94, Pasadena, CA 91125, U. S. A.

(Received Oct. 28, 2022 / Revised Feb. 3, 2023)

Abstract We present a Gaussian process (GP) approach, called Gaussian process hydrodynamics (GPH) for approximating the solution to the Euler and Navier-Stokes (NS) equations. Similar to smoothed particle hydrodynamics (SPH), GPH is a Lagrangian particle-based approach that involves the tracking of a finite number of particles transported by a flow. However, these particles do not represent mollified particles of matter but carry discrete/partial information about the continuous flow. Closure is achieved by placing a divergence-free GP prior ξ on the velocity field and conditioning it on the vorticity at the particle locations. Known physics (e.g., the Richardson cascade and velocity-increment power laws) is incorporated into the GP prior by using physics-informed additive kernels. This is equivalent to expressing ξ as a sum of independent GPs ξ^l, which we call modes, acting at different scales (each mode ξ^l self-activates to represent the formation of eddies at the corresponding scales). This approach enables a quantitative analysis of the Richardson cascade through the analysis of the activation of these modes, and enables us to analyze coarse-grain turbulence statistically rather than deterministically. Because GPH is formulated by using the vorticity equations, it does not require solving a pressure equation. By enforcing incompressibility and fluid-structure boundary conditions through the selection of a kernel, GPH requires significantly fewer particles than SPH. Because GPH has a natural probabilistic interpretation, the numerical results come with uncertainty estimates, enabling their incorporation into an uncertainty quantification (UQ) pipeline and adding/removing particles (quanta of information) in an adapted manner. The proposed approach is suitable for analysis because it inherits the complexity of state-of-the-art solvers for dense kernel matrices and results in a natural definition of turbulence as information loss. Numerical experiments support the importance of selecting physics-informed kernels and illustrate the major impact of such kernels on the accuracy and stability. Because the proposed approach uses a Bayesian interpretation, it naturally enables data assimilation and predictions and estimations by mixing simulation data and experimental data.

Key words Navier-Stokes (NS) equation, Euler, Lagrangian, vorticity, Gaussian process (GP), physics-informed kernel

Chinese Library Classification O241.8, O357.1, O357.4
2010 Mathematics Subject Classification 35Q30, 76D05, 60G15, 65M75, 65N75, 65N35, 47B34, 41A15, 34B15

* Citation: OWHADI, H. Gaussian process hydrodynamics. *Applied Mathematics and Mechanics* (*English Edition*), **44**(7), 1175–1198 (2023) https://doi.org/10.1007/s10483-023-2990-9
† Corresponding author, E-mail: owhadi@caltech.edu

1 Introduction

The Navier-Stokes (NS) equations are difficult to both analyze[1] and approximate numerically because of the emergence of multiple nonlinearly coupled scales. Even from a physicist's perspective, they remain poorly understood, and we still do not have a clear definition of turbulence beyond "the complex, chaotic motion of a fluid"[2]. The NS equations are also difficult to solve because they contain a dual description of the underlying physics that is Lagrangian in its representation of Newton's second law and Eulerian in its description of the pressure equation. Thus, classical methods for solving the NS equations are divided into Eulerian (grid-based) and Lagrangian (meshfree particle-based) methods. While Eulerian methods are more efficient in solving pressure equations, they require a high resolution to solve the Lagrangian effects of the equations. While Lagrangian methods are efficient at replicating conservation laws (e.g., entropy, momentum, and energy), they require a large number of particles to solve the Eulerian aspects of the equations (e.g., solve for the pressure given the position/velocities of the particles).

1.1 Smoothed particle hydrodynamics (SPH)

SPH is a prototypical Lagrangian meshfree particle method (in which the continuum is assumed to be a collection of imaginary particles) introduced in the late 1970s for astrophysics problems[3–4] (see Ref. [5] for a review). Although SPH has been widely applied to various areas of engineering and science (see Ref. [6] for an overview), including computational fluid dynamics (CFD), it suffers from the difficulties associated with the Lagrangian methods and "still requires development to address important elements which prevent more widespread use"[7]. These elements (identified as major challenges in Ref. [7]) include (i) convergence, consistency, and stability, (ii) boundary conditions, (iii) adaptivity, (iv) coupling to other models, and (v) applicability to the industry.

1.2 Gaussian process hydrodynamics (GPH)

This paper introduces GPH as an information/inference-based approach into approximating the NS equations. Although numerical approximations and statistical inferences may be considered as separate subjects, they are intimately connected through the common objective of forming estimations with partial information[8], and kernel/GP methods provide a natural (and minimax optimal[9]) approach to computing with missing information. In the proposed GPH approach, flow-advected particles carry partial information regarding the underlying vorticity/velocity fields, and (information gap) closure is achieved by randomizing the underlying velocity field via a Gaussian process (GP) prior with a physics-informed kernel, ensuring that incompressibility and boundary conditions are exactly satisfied and power/scaling and energy transfer laws are satisfied statistically. From this perspective, turbulence can be defined and quantified as information loss between the true dynamics of the NS equations and those resulting from carrying only partial information about the underlying fields. Although GPH has similarities with SPH, it also has several significant differences. (i) In SPH, particles represent mollified particles of matter, whereas in GPH, particles represent discrete/partial information about the continuous flow. (ii) SPH is typically formulated on the velocity and requires solving a pressure equation, whereas GPH is formulated on the vorticity equations and Eulerian aspects (e.g., recovering the velocity field) are solved by using GP regression. (iii) By enforcing incompressibility and fluid-structure boundary conditions through the selection of the kernel, GPH requires significantly fewer particles. (iv) By carrying variance information, GPH enables the adding and removing of quanta of information from the flow in an adapted manner. While SPH recovers fields through smooth approximations of delta Dirac functions with compactly supported kernels, GPH focuses on the optimal recovery[9–10] of the missing information with adapted/programmed kernels[11]. Its representation of the multiscale structure of the flow

through regression additive kernels enables a corresponding statistical decomposition of the flow at different scales (modes), and a quantitative analysis of the Richardson cascade through the analysis of the activation of these modes[11]. Its focus on informing the kernel about the underlying physics and boundary conditions creates a different strategy for solving some of the major challenges of SPH listed above. Its probabilistic/Bayesian interpretation enables it to be incorporated into uncertainty quantification (UQ) pipelines.

1.3 Vortex methods

Because GPH resembles vortex methods[12–13] (owing to its formulation on the vorticity equations), it can also be interpreted as a generalization of such methods to arbitrary kernel approximations of the underlying vorticity and velocity fields based on discrete vorticity information carried by the Lagrangian particles. However, the velocity field is not recovered from the continuous vorticity field using the Biot-Savart law but from the available partial information about the continuous vorticity field using kernel (GP regression) representer formulae.

1.4 Solving partial differential equations (PDEs) as learning problems

Two main approaches are available for solving PDEs as learning problems: (i) artificial neural network (ANN)-based approaches, with physics-informed neural networks[14–15] as a prototypical example and (ii) GP-based approaches, with Gamblets[16–18] as a prototypical example. Although GP-based approaches are more theoretically well-founded[9] and have a long history of interplay with numerical approximation[8,19–21], they were essentially limited to linear/quasi-linear/time-dependent PDEs and have only recently been generalized to arbitrary nonlinear PDEs[22] (and computational graphs[23]).

1.5 Physics-informed kernels

While both the ANN and GP methods replace the solution to the PDE with an ANN/GP and are physics-informed by constraining/conditioning the ANN/GP to satisfy the PDE over a finite number of degrees of freedom (e.g., collocation points), GP methods can also be physics-informed through their kernels[16]. The importance of employing physics/PDE-informed kernels is well understood in numerical approximation/homogenization by using Darcy's elliptic PDE $-\operatorname{div}(a\nabla)$ (with a rough conductivity a) as a prototypical example. While employing a smooth kernel may result in arbitrary poor convergence[24], employing a physics-informed kernel ensures an optimal convergence rate[16]. Although Owhadi[16] proposed identifying such kernels by filtering white noise using the solution operator of the PDE (i.e., replacing the right-hand side/source term with white noise and conditioning the resulting randomized solution on a finite number of linear measurements), this approach is impractical for nonlinear PDEs because the resulting solution is not a GP.

The approach proposed in this paper involves selecting a physics-informed kernel by programming the kernel[11] to satisfy (i) the divergence-free condition of the velocity field, (ii) the boundary conditions, (iii) the statistical power laws, and (iv) the Richardson cascade of turbulence.

1.6 Outline of the article

The remainder of the article is organized as follows. Sections 2 and 3 introduce GPH in the setting of the vorticity formulation of the forced NS equations. Section 4 describes representer formulae for the underlying GP formulation with divergence-free kernels. Section 5 describes the design of physics-informed kernels for GPH. Section 6 quantifies the accuracy of the proposed approach as the L^2 norm of its residual, interprets that residual as an instantaneous measure of information loss (resulting from the discretization of continuous dynamics), and presents an information loss interpretation and quantification of turbulence. Section 7 presents numerical experiments. In all these sections, we use figures and simulations from Section 7 to illustrate the proposed method. Please refer to Section 7 for their detailed descriptions and to https://www.youtube.com/user/HoumanOwhadi for corresponding animations. Section 8 presents further discussion.

2 Set up

Let $\mathbb{T}^d$ be a torus of side length 2π and dimension $d = 2$ or 3. Consider the forced NS equations on $\mathbb{T}^d$ as

$$\begin{cases} \partial_t u + u\nabla u = \nu\Delta u - \nabla p + f \quad \text{on} \quad \mathbb{T}^d, \\ \operatorname{div} u = 0 \quad \text{on} \quad \mathbb{T}^d \end{cases} \tag{1}$$

with a smooth zero-mean flow initial condition $u(x,0) = u_0(x)$ and an external volumetric force $f(x,t)$, where $u_0 \in C^\infty(\mathbb{T}^d)$, $\int_{\mathbb{T}^d} u_0(x)\,\mathrm{d}x = 0$, $f \in C^\infty(\mathbb{T}^d \times [0,\infty))$, and $\int_{\mathbb{T}^d} f(x,t)\,\mathrm{d}x = 0$ for all values of t.

By introducing the vorticity

$$\omega(x,t) := \operatorname{curl} u(x,t) \tag{2}$$

and $g(x,t) = \operatorname{curl} f(x,t)$, (1) is equivalent to the following equations:

$$\partial_t\omega + u\nabla\omega = \nu\Delta\omega + g(x,t), \quad d = 2, \tag{3}$$

$$\partial_t\omega + u\nabla\omega = \nu\Delta\omega + \omega\nabla u + g(x,t), \quad d = 3 \tag{4}$$

with the initial condition $\omega(x,0) = \omega_0(x) := \operatorname{curl} u_0(x)$.

3 GPH

Let $X_1, X_2, \cdots, X_N$ be N distinct (and possibly homogeneously distributed) collocation points in $\mathbb{T}^d$. For $i \in \{1,2,\cdots,N\}$, let $t \to q_i(t)$ be the trajectory formed by a particle advected by the flow velocity $u(x,t)$, which is defined as the solution to

$$\dot{q}_i(t) = u(q_i(t),t) \tag{5}$$

with the initial condition $q_i^0 = X_i \in \mathbb{T}^d$. For $i \in \{1,2,\cdots,N\}$, let

$$W_i(t) := \omega(q_i(t),t) \tag{6}$$

be the value of the vorticity at $(q_i(t),t)$. (5), (3), and (4) imply that $t \to W_i(t)$ solves the ordinary differential equations (ODEs)

$$\dot{W}_i(t) = \nu\Delta\omega(q_i(t),t) + g(q_i(t),t), \quad d = 2, \tag{7}$$

$$\dot{W}_i(t) = \nu\Delta\omega(q_i(t),t) + W_i(t)\nabla u(q_i(t),t) + g(q_i(t),t), \quad d = 3 \tag{8}$$

with the initial condition $W_i(0) = \omega_0(q_i(0))$. Write $q(t) := (q_1(t), q_2(t), \cdots, q_N(t))$ and $W(t) := (W_1(t), W_2(t), \cdots, W_N(t))$. Because (q,W) provides only partial information on u and its partial derivatives, (7) and (8) are not autonomous systems, and closing them requires closing the information gap between (q,W) and u, i.e., approximating $u(x,t)$ and its partial derivatives as a function of (q,W). Our approach to this closure problem involves replacing the unknown velocity field u by a centered GP $\xi \sim \mathcal{N}(0,K)$ (with a physics-informed matrix-valued kernel K that may be non-stationary to incorporate non-periodic boundary conditions) and approximating u with the conditional expectation of ξ given the information (6). To describe this, let

$$\mathcal{Y} := (\mathbb{T}^2)^N \times \mathbb{R}^N, \quad d = 2, \tag{9}$$

$$\mathcal{Y} := (\mathbb{T}^3)^N \times (\mathbb{R}^3)^N, \quad d = 3 \tag{10}$$

be the phase space containing the trajectory $t \to (q, W)(t)$. Define

$$u^\star(x, q, W) := \mathbb{E}(\xi(x) \big| \operatorname{curl} \xi(q) = W) \quad \text{for} \quad (x, q, W) \in \mathbb{T}^d \times \mathcal{Y}. \tag{11}$$

With vectorized notations, we can write $\operatorname{curl} \xi(q)$ for the N-vector with entries $\operatorname{curl} \xi(q_i)$. We then approximate $(q, W)(t)$ with $(q^\star, W^\star)(t)$, $u(x, t)$ with

$$\overline{u}(x, t) := u^\star(x, q^\star(t), W^\star(t)), \tag{12}$$

and $\omega(x, t)$ with

$$\overline{\omega}(x, t) := \operatorname{curl} u^\star(x, q^\star(t), W^\star(t)), \tag{13}$$

where $(q^\star, W^\star)$ is the solution to the autonomous system of ODEs (the differential operators $\Delta \operatorname{curl}$ and ∇ in (14) act on the first argument x of $u^\star$ in (11)),

$$\begin{cases} \dot{q}_i^\star = u^\star(q_i^\star, q^\star, W^\star), \\ \dot{W}_i^\star(t) = \nu \Delta \operatorname{curl} u^\star(q_i^\star, q^\star, W^\star) + g(q_i^\star(t), t), \quad d = 2, \\ \dot{W}_i^\star(t) = \nu \Delta \operatorname{curl} u^\star(q_i^\star, q^\star, W^\star) + W_i^\star \nabla u^\star(q_i^\star, q^\star, W^\star) + g(q_i^\star(t), t), \quad d = 3 \end{cases} \tag{14}$$

with the initial condition $(q^\star, W^\star)(0) = (q, W)(0) = (q^0, \omega_0(q^0))$. Figures 1 and 2 are the snapshots of $\overline{u}$ (shown as a vector field), $\overline{\omega}$ (shown as a heatmap), and $\overline{q}$ (shown as dark points) of a point $x = (x_1, x_2)$ on the two-dimensional (2D) torus. Figure 3 shows the snapshots of $\overline{u}$ (shown as blue arrows), W (shown as red arrows), and $\overline{q}$ (shown as dark points) of a point $x = (x_1, x_2, x_3)$ on the three-dimensional (3D) torus for $d = 3$.

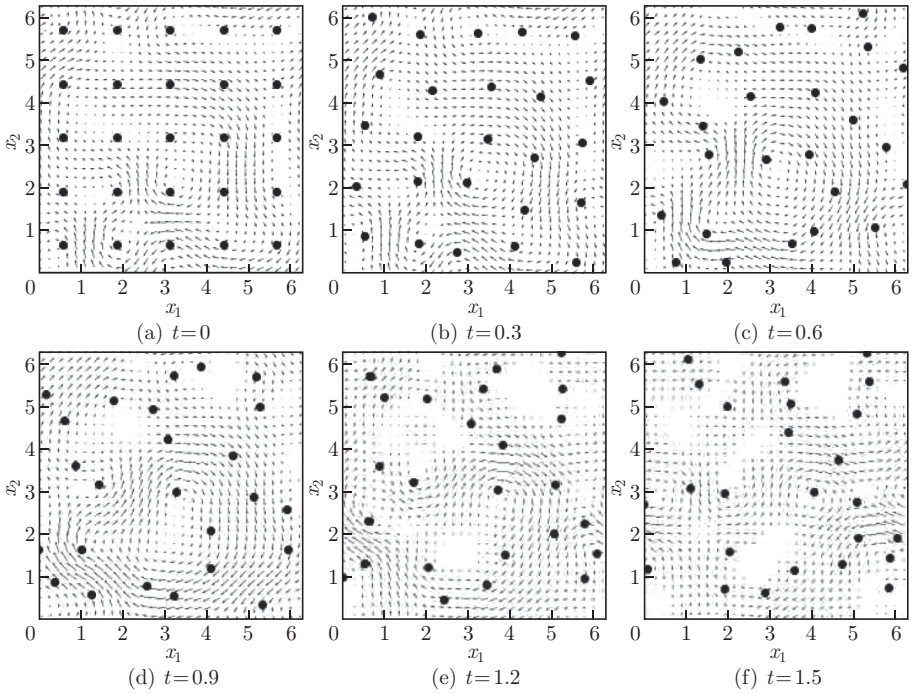

Fig. 1 Velocity snapshots

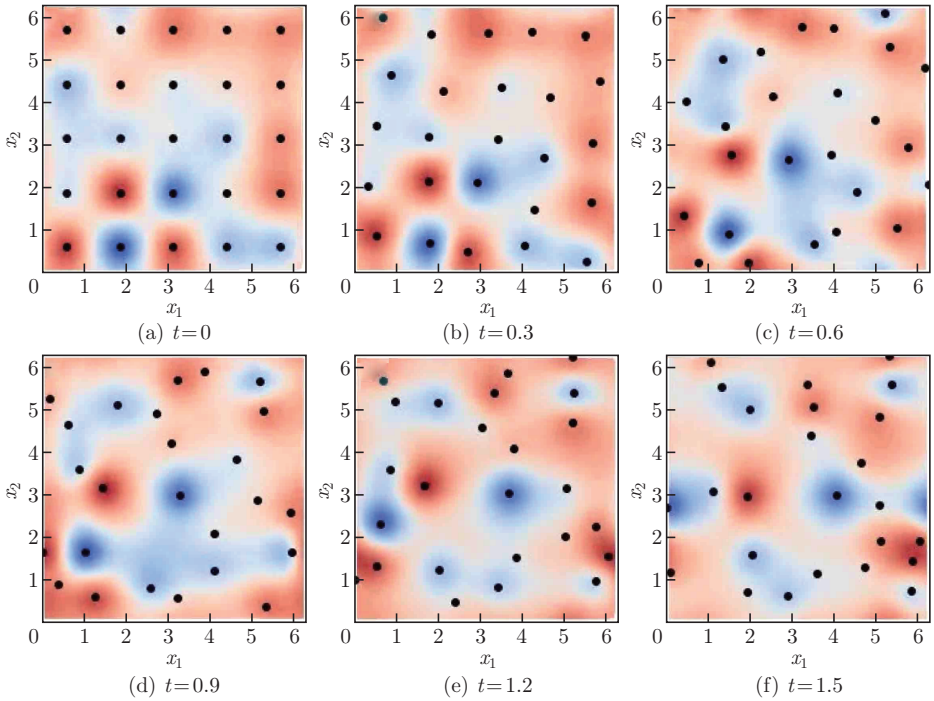

Fig. 2 Vorticity snapshots (color online)

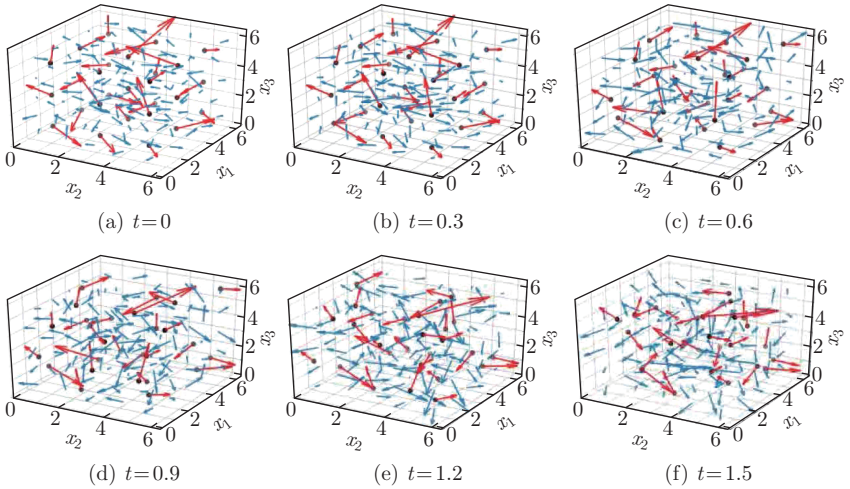

Fig. 3 Velocity and vorticity snapshots for $d = 3$, where the blue and red arrows show the velocity and the vorticity W at particle locations q, respectively (color online)

The proposed approach contains UQ estimates and is compatible with a UQ pipeline. In particular, given curl $\xi(q) = W$, ξ is a GP with conditional mean $u^\star$ and conditional covariance function,

$$\mathcal{C}^u(x,y) := \mathbb{E}((\xi(x) - u^\star(x,q,W))(\xi(y) - u^\star(x,q,W))^{\mathrm{T}} | \operatorname{curl}\xi(q) = W), \tag{15}$$

and curl ξ is a GP with conditional mean curl $u^\star$ and conditional covariance function

$$\mathcal{C}^\omega(x,y) := \mathbb{E}((\operatorname{curl}\xi(x) - \operatorname{curl}u^\star(x,q,W))(\operatorname{curl}\xi(y) - \operatorname{curl}u^\star(x,q,W))^{\mathrm{T}} | \operatorname{curl}\xi(q) = W). \tag{16}$$

Figures 4 and 5 present the snapshots of $x \to \mathrm{Tr}(\mathcal{C}^u(x,x))$ and the heatmaps of $x \to \mathrm{Tr}(\mathcal{C}^\omega(x,x))$, respectively.

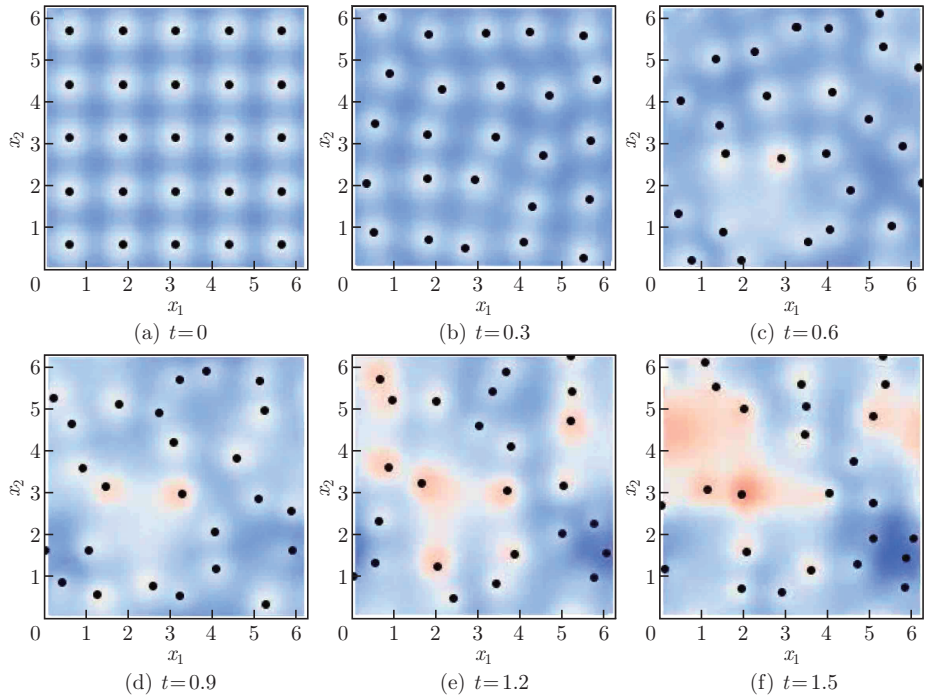

Fig. 4 Variance velocity, where the color scale ranges from 4.8 (blue) to 8.4 (red) (color online)

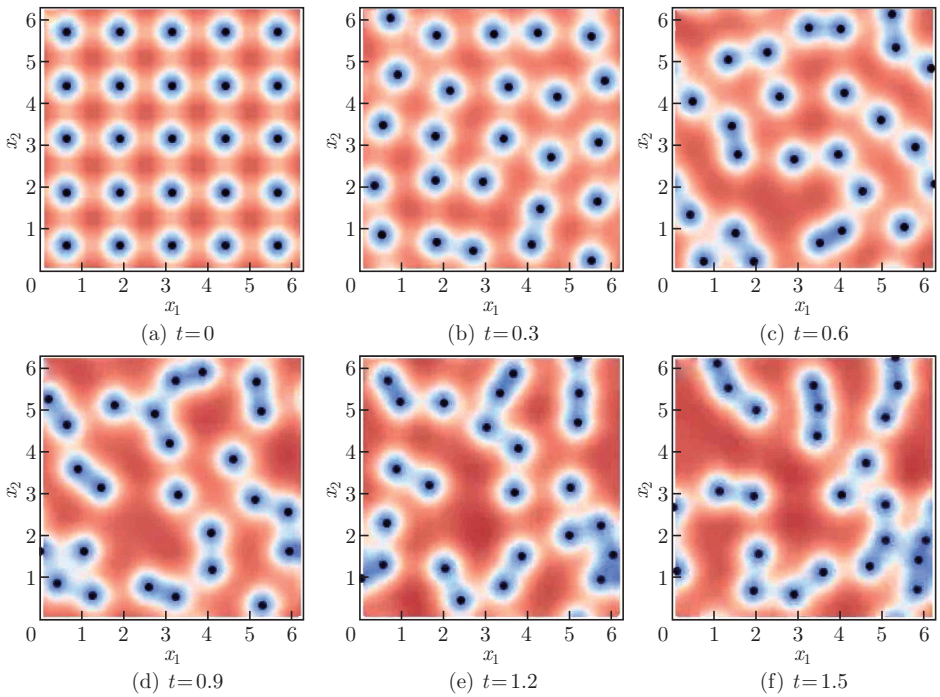

Fig. 5 Variance vorticity, where the color scale ranges from 0 (blue) to 45 (red) (color online)

4 Divergence free GPs/kernels and representer formulae

We now describe the vector-valued GP $\xi \sim \mathcal{N}(0, K)$ to close the NS equations and introduce the representer formulae for identifying $u^\star$ and its partial derivatives as a function of $(q^\star, W^\star)$. Recall (see Chapters 7 and 17 in Ref. [9] and Subsection 8.1 in Ref. [25]) that $x \to \xi(x)$ is a map from $\mathbb{T}^d$ to a linear (Hilbert) space of d-dimensional centered Gaussian vectors such that

$$\mathrm{Cov}(\xi(x), \xi(y)) = K(x, y) \quad \text{for} \quad x, y \in \mathbb{T}^d, \tag{17}$$

where the covariance function K is a $\mathbb{R}^{d\times d}$ matrix-valued kernel (also known as a vector-valued kernel[26]). Write $\mathcal{H}_K$ for the reproducing kernel Hilbert space (RKHS) of $\mathbb{R}^d$ valued functions defined by K. To ensure that our approximation $u^\star$ remains zero-mean and incompressible, and that (11) and (14) are properly defined, we select K such that $\mathcal{H}_K$ is contained in the set

$$\mathcal{S}^3(\mathbb{T}^d) := \Big(v \in C^3(\mathbb{T}^d) \mid \int_{\mathbb{T}^d} v(x)\,\mathrm{d}x = 0 \quad \text{and} \quad \mathrm{div}\, v = 0\Big) \tag{18}$$

of $\mathbb{T}^d$-periodic zero-mean divergence-free $\mathbb{R}^d$-valued functions with continuous third-order derivatives (we write C^k for the space of continuously kth-order differentiable functions). Matrix-valued kernels inducing an RKHS containing divergence-free vector-valued functions can be constructed by starting with a stationary scalar-valued kernel $G(x, x') = g(x - x')$ and selecting $K(x, y) = (\mathrm{Hess}\, g - \mathrm{Tr}(\mathrm{Hess}\, g)I_d)(x - y)$, where Hess is the Hessian operator, and I_d is the $d \times d$ identity matrix (see Subsection 5.1 in Ref. [26]). Here, we present a more general approach enabling using non-stationary kernels and the incorporation of nontrivial fluid-structure boundary conditions into the kernel (see Fig. 6). We will distinguish between $d = 2$ and $d = 3$ cases in our description of this approach.

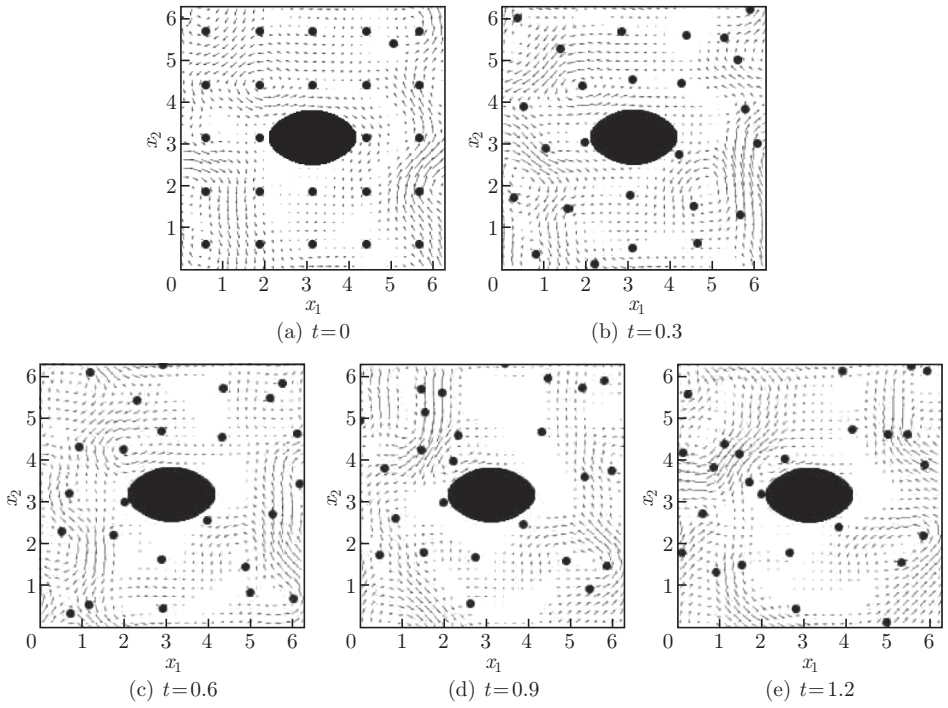

Fig. 6 Flow around an obstacle

4.1 2D case

4.1.1 Divergence-free kernels

Given an $\mathbb{R}^2$-valued function $v(x) = (v_1(x), v_2(x))^{\mathrm{T}}$, $\operatorname{curl} v = (-\partial_{x_2} v_1(x) + \partial_{x_1} v_2(x))$ can be written as the inner product between the row vector $\operatorname{curl}_x = (-\partial_{x_2}, \partial_{x_1})$ and the column vector $v(x) = (v_1(x), v_2(x))^{\mathrm{T}}$. Let G be a non-degenerate C^3-differentiable scalar-valued kernel on $\mathbb{T}^2$ such that $\mathcal{H}_G$ (the RKHS defined by G) is compactly embedded in $H^s(\mathbb{T}^2)$ for $s > 5$. Extending matrix-vector operations to differential operators, we define

$$K(x,y) := \operatorname{curl}_x^{\mathrm{T}} \operatorname{curl}_y G(x,y) = \begin{pmatrix} -\partial_{x_2} \\ \partial_{x_1} \end{pmatrix} \begin{pmatrix} -\partial_{y_2} & \partial_{y_1} \end{pmatrix} G(x,y), \tag{19}$$

which can also be written as

$$K(x,y) := \begin{pmatrix} \partial_{x_2}\partial_{y_2} & -\partial_{x_2}\partial_{y_1} \\ -\partial_{x_1}\partial_{y_2} & \partial_{x_1}\partial_{y_1} \end{pmatrix} G(x,y) := \begin{pmatrix} \partial_{x_2}\partial_{y_2} G(x,y) & -\partial_{x_2}\partial_{y_1} G(x,y) \\ -\partial_{x_1}\partial_{y_2} G(x,y) & \partial_{x_1}\partial_{y_1} G(x,y) \end{pmatrix}. \tag{20}$$

The following proposition shows that K is a valid non-degenerate kernel satisfying our requirements.

Proposition 1 *It holds true that* (i) $K =$(20) *is a non-degenerate kernel*, (ii) *its RKHS* $\mathcal{H}_K$ *is compactly embedded in* $\mathcal{H}^{s-1}(\mathbb{T}^2)$, *and* (iii) $\mathcal{H}_K \subset \mathcal{S}^3(\mathbb{T}^2)$.

Proof To show that K is a valid kernel, we will employ the one-to-one map among kernels, symmetric positive definite linear operators, and quadratic norms presented in Chapters 11 and 17 in Ref. [9] (see also Subsection 2.1 in Ref. [22]). Define $\mathcal{H}_G$ and $\|\cdot\|_G$ as the RKHS space and the RKHS norm induced by G and $\mathcal{H}_G^*$, and $\|\cdot\|_G^*$ as their duals with respect to the L^2 inner product which we express as $[\cdot,\cdot]$ ($[\varphi, f] := \int_{\mathbb{T}^2} \varphi(x) f(x)\,\mathrm{d}x$ for $\varphi \in \mathcal{H}_G^*$ and $f \in \mathcal{H}_G$). The operation $\varphi \to \int_{\mathbb{T}^2} G(x,y)\varphi(y)\,\mathrm{d}y$ defines a linear bijection $\mathcal{G}$ mapping $\mathcal{H}_G^*$ to $\mathcal{H}_G$ that is symmetric ($[\varphi, \mathcal{G}\varphi'] = [\varphi', \mathcal{G}\varphi]$), positive ($[\varphi, \mathcal{G}\varphi] \geqslant 0$), and definite ($[\varphi, \mathcal{G}\varphi] = 0$ if and only if $\varphi = 0$). With δ_x as a delta Dirac function supported at the point x, $\mathcal{G}$ defines the kernel G via $G(x,y) = [\delta_x, \mathcal{G}\delta_y]$. Furthermore, $\|\varphi\|_G^{*,2} = \int_{(\mathbb{T}^2)^2} G(x,y)\varphi(x)\varphi(y)\,\mathrm{d}x\,\mathrm{d}y = [\varphi, \mathcal{G}\varphi]$ for $\varphi \in \mathcal{H}_G$, and $\|f\|_G = \sup_{\varphi \in \mathcal{H}_G^*} [\varphi, f]/\|\varphi\|_G^*$ for $f \in \mathcal{H}_G$. These identities indicate that there is a one-to-one correspondence among the (non-degenerate) kernel G, the symmetric positive definite linear bijection $\mathcal{G}$, and the quadratic norms $\|\cdot\|_G^*$ and $\|\cdot\|_G$ (any of these objects can be used to define a valid kernel, as shown in Chapters 11 and 17 in Ref. [9]). For $\phi \in \mathcal{S}^3(\mathbb{T}^2)$, write

$$\|\phi\|_K^* := \|\operatorname{curl}\phi\|_G^*. \tag{21}$$

Because $\|\phi\|_K^*$ is a quadratic norm on $\mathcal{S}^3(\mathbb{T}^2)$, it defines a non-degenerate kernel K (see Chapters 11 and 17 in Ref. [9]) with the RKHS space $\mathcal{H}_K$ and norm $\|\cdot\|_K$ such that $\|\cdot\|_K^*$ is the dual of $\|\cdot\|_K$ with respect to the L^2 inner product and $\mathcal{H}_K^*$ is the closure of $\mathcal{S}^3(\mathbb{T}^2)$ with respect to $\|\cdot\|_K^*$ (note that the construction (19) and the identity $\operatorname{div}\operatorname{curl} = \mathbf{0}$ imply that the elements of $\mathcal{H}_K$ are divergence-free functions). For the sake of clarity, we will also present the following alternative proof of the non-degeneracy of K. For $q \in (\mathbb{T}^2)^N$, $K(q,q)$ is the $N \times N$ block matrix with 2×2 block entries $K(q_i, q_j)$. For $\alpha \in (\mathbb{R}^2)^N$, write $\alpha^{\mathrm{T}} K(q,q)\alpha := \sum_{i,j=1}^{N} \alpha_i^{\mathrm{T}} K(q_i, q_j)\alpha_j$. Thus, the identity $\alpha^{\mathrm{T}} K(q,q)\alpha = \big\| \sum_{i=1}^{N} \delta_{q_i} \circ (-\alpha_1 \partial_{x_2} + \alpha_2 \partial_{x_1}) \big\|_G^{*,2}$ implies that $K(q,q)$ is invertible if q_i are pairwise distinct and $\alpha \neq 0$, i.e., K is non-degenerate. (ii) in Proposition 1 follows from the identity (21) and $\|v\|_K = \sup_{\phi \in \mathcal{H}_K^*} [\phi, v]/\|\phi\|_K^*$. (iii) in Proposition 1 follows from the compact embedding of $\mathcal{H}_G$ into $H^s(\mathbb{T}^2)$ for $s > 5$, and the compact embedding of $H^{s-1}(\mathbb{T}^2)$ into $C^3(\mathbb{T}^2)$ for $s > 5$.

Remark 1 The results of this section can naturally be generalized to the scenario in which G is a kernel on A/B, where B is an inclusion in the domain A. In this case, the required boundary conditions on the elements of $\mathcal{H}_K$ (e.g., stick or no-slip) transfer onto the required boundary conditions on G. Possible designs of G include (i) identifying G as Green's function of a higher-order elliptic PDE on A/B with the required boundary conditions and (ii) designing G with the transformations of an initial defined on $\mathbb{R}^d$. For Fig. 6, $G(x,y) = \mathfrak{g}(x,y)f(x)f(y)$ where $\mathfrak{g}$ is a kernel on $\mathbb{T}^2$, and f is a smooth function equal to 0 when B is included and 1 when $\mathbb{T}^2/B_\epsilon$ is included, where B_ϵ is an ϵ enlargement of B obtained by adding a boundary layer of size ϵ (the resulting elements of $\mathcal{H}_K$ satisfy a stick boundary condition).

4.1.2 Representer formulae

We now introduce the representer formulae for the conditional mean and covariance of the GP $\xi \sim \mathcal{N}(0,K)$ given $\operatorname{curl}\xi(q) = W$. $\Delta_x\Delta_y G$ represents the $\mathbb{T}^2$ valued kernel $\Delta_x\Delta_y G(x,y)$. For $q \in (\mathbb{T}^2)^N$, $\Delta_x\Delta_y G(q,q)$ represents the $N \times N$ matrix with entries $\Delta_x\Delta_y G(q_i,q_j)$. For $x \in \mathbb{T}^2$ and $q \in \mathbb{T}^N$, $\operatorname{curl}_x^{\mathrm{T}}\Delta_y G(x,q)$ represents the N-vector with $\mathbb{R}^2$-valued entries $\operatorname{curl}_x^{\mathrm{T}}\Delta_y G(x,q_i)$.

Proposition 2 *The GP $\xi \sim \mathcal{N}(0,K)$ conditioned on $\operatorname{curl}\xi(q) = W$ is also Gaussian with the conditional mean*

$$u^\star(x,q,W) = (11) = \operatorname{curl}_x^{\mathrm{T}}\Delta_y G(x,q)(\Delta_x\Delta_y G(q,q))^{-1}W \tag{22}$$

and the conditional covariance kernel $\mathcal{C}^u(x,y) =$(15) given by

$$\operatorname{curl}_x^{\mathrm{T}}\operatorname{curl}_y G(x,y) - \operatorname{curl}_x^{\mathrm{T}}\Delta_y G(x,q)(\Delta_x\Delta_y G(q,q))^{-1}\Delta_x\operatorname{curl}_y G(q,y). \tag{23}$$

Furthermore, the GP curl ξ *conditioned on* $\operatorname{curl}\xi(q) = W$ *is also Gaussian with the conditional mean*

$$\operatorname{curl}u^\star(x,q,W) = \Delta_x\Delta_y G(x,q)(\Delta_x\Delta_y G(q,q))^{-1}W \tag{24}$$

and the conditional covariance kernel $\mathcal{C}^\omega(x,y) =$(16) given by

$$\Delta_x\Delta_y G(x,y) - \Delta_x\Delta_y G(x,q)(\Delta_x\Delta_y G(q,q))^{-1}\Delta_x\Delta_y G(q,y). \tag{25}$$

Proof (22) and (23) follow from the generalized representer theorem (see Corollary 17.12 in Ref. [9] and Proposition 2.1 in Ref. [22]) and the identity $\operatorname{curl}\operatorname{curl}^{\mathrm{T}} = \Delta$. For $\alpha \in \mathbb{R}^N$, $\alpha^{\mathrm{T}}\Delta_x\Delta_y G(q,q)\alpha = \|\sum_{i=1}^N \alpha_i\delta_{q_i}\circ\Delta\|_G^{*,2}$ implies that $\Delta_x\Delta_y G(q,q)$ is invertible if q_i are pairwise distinct. (24) and the identities $\operatorname{curl}\operatorname{curl}^{\mathrm{T}} = \Delta$ and $\mathcal{C}^\omega(x,y) = \operatorname{curl}_x\operatorname{curl}_y^{\mathrm{T}}\mathcal{C}^u(x,y)$ imply (24) and (25).

Using Proposition 2, (14) reduces to

$$\begin{cases} \dot{q}_i^\star = \operatorname{curl}_x^{\mathrm{T}}\Delta_y G(q_i^\star,q^\star)(\Delta_x\Delta_y G(q^\star,q^\star))^{-1}W^\star, \\ \dot{W}_i^\star = \nu\Delta_x^2\Delta_y G(q_i^\star,q^\star)(\Delta_x\Delta_y G(q^\star,q^\star))^{-1}W^\star + g(q_i^\star(t),t). \end{cases} \tag{26}$$

4.2 3D case

4.2.1 Divergence-free kernels

Given an $\mathbb{R}^3$-valued function $v(x)$, $\operatorname{curl}v$ can be written as the inner product between the matrix

$$\operatorname{curl}_x = \begin{pmatrix} 0 & -\partial_{x_3} & \partial_{x_2} \\ \partial_{x_3} & 0 & -\partial_{x_1} \\ -\partial_{x_2} & \partial_{x_1} & 0 \end{pmatrix} \tag{27}$$

and the column vector $v(x) = (v_1(x), v_2(x), v_3(x))^{\mathrm{T}}$. Let G be a non-degenerate C^3-differentiable scalar-valued kernel on $\mathbb{T}^3$ such that $\mathcal{H}_G$ (the RKHS defined by G) is compactly embedded in $H^s(\mathbb{T}^3)$ for $s > 5.5$. Define

$$K(x,y) := \operatorname{curl}_x^{\mathrm{T}} \operatorname{curl}_y G(x,y) = \begin{pmatrix} 0 & \partial_{x_3} & -\partial_{x_2} \\ -\partial_{x_3} & 0 & \partial_{x_1} \\ \partial_{x_2} & -\partial_{x_1} & 0 \end{pmatrix} \begin{pmatrix} 0 & -\partial_{y_3} & \partial_{y_2} \\ \partial_{y_3} & 0 & -\partial_{y_1} \\ -\partial_{y_2} & \partial_{y_1} & 0 \end{pmatrix} G(x,y),$$

which can also be written as

$$K(x,y) := \begin{pmatrix} \partial_{x_3}\partial_{y_3} + \partial_{x_2}\partial_{y_2} & -\partial_{x_2}\partial_{y_1} & -\partial_{x_3}\partial_{y_1} \\ -\partial_{x_1}\partial_{y_2} & \partial_{x_3}\partial_{y_3} + \partial_{x_1}\partial_{y_1} & -\partial_{x_3}\partial_{y_2} \\ -\partial_{x_1}\partial_{y_3} & -\partial_{x_2}\partial_{y_3} & \partial_{x_1}\partial_{y_2} + \partial_{x_2}\partial_{y_2} \end{pmatrix} G(x,y). \tag{28}$$

Proposition 3 *It holds true that* (i) $K = (28)$ *is a non-degenerate kernel*, (ii) *its RKHS* $\mathcal{H}_K$ *is compactly embedded in* $\mathcal{H}^{s-1}(\mathbb{T}^3)$, *and* (iii) $\mathcal{H}_K \subset \mathcal{S}^3(\mathbb{T}^3)$.

Proof With Hess as the Hessian operator and I_3 as the 3×3 identity matrix, integrating by parts, we observe that for $\phi \in \mathcal{S}^3(\mathbb{T}^3)$,

$$\|\phi\|_K^{*,2} = \| - \partial_{x_3}\phi_2 + \partial_{x_2}\phi_1\|_G^{*,2} + \|\partial_{x_3}\phi_1 - \partial_{x_1}\phi_3\|_G^{*,2} + \| - \partial_{x_2}\phi_1 + \partial_{x_1}\phi_2\|_G^{*,2}. \tag{29}$$

The remainder of the proof is identical to that of Proposition 1.

4.2.2 Representer formulae

We now present representer formulae for the conditional mean and covariance of the GP $\xi \sim \mathcal{N}(0, K)$ given $\operatorname{curl}\xi(q) = W$. Define $\mathcal{L}_x := I_3\Delta_x - \operatorname{Hess}_x$ and $\mathcal{L}_x\mathcal{L}_y G$ as the 3×3 matrix valued kernel obtained by letting $\mathcal{L}_x$ act on the x variable and $\mathcal{L}_y$ act on the y variable of $G(x,y)$. Similarly, we define $\operatorname{curl}_x^{\mathrm{T}} \mathcal{L}_y G$ as the 3×3 matrix valued function of x and y, which is obtained by letting $\operatorname{curl}_x^{\mathrm{T}}$ act on the x variable and $\mathcal{L}_y$ on the y variable of $G(x,y)$. Using the shorthand notations of Subsection 4.1.2, for $q \in (\mathbb{T}^3)^N$, we define $\mathcal{L}_x\mathcal{L}_y G(q,q)$ as the $N \times N$ block matrix whose entries are the 3×3 matrices $\mathcal{L}_x\mathcal{L}_y G(q_i, q_j)$. Similarly, we define $\operatorname{curl}_x^{\mathrm{T}} \mathcal{L}_y G(x,q)$ as the N-block vector whose entries are the 3×3 matrices $\operatorname{curl}_x^{\mathrm{T}} \mathcal{L}_y G(x,q_i)$.

Proposition 4 *The GP* $\xi \sim \mathcal{N}(0,K)$ *conditioned on* $\operatorname{curl}\xi(q) = W$ *is also Gaussian with the conditional mean*

$$u^\star(x,q,W) = (11) = \operatorname{curl}_x^{\mathrm{T}} \mathcal{L}_y G(x,q)(\mathcal{L}_x\mathcal{L}_y G(q,q))^{-1} W \tag{30}$$

and the conditional covariance kernel $\mathcal{C}^u(x,y) = (15)$ *given by*

$$\operatorname{curl}_x^{\mathrm{T}} \operatorname{curl}_y G(x,y) - \operatorname{curl}_x^{\mathrm{T}} \mathcal{L}_y G(x,q)(\mathcal{L}_x\mathcal{L}_y G(q,q))^{-1} \mathcal{L}_x \operatorname{curl}_y G(q,y). \tag{31}$$

Furthermore, the GP $\operatorname{curl}\xi$ *conditioned on* $\operatorname{curl}\xi(q) = W$ *is also Gaussian with the conditional mean*

$$\operatorname{curl} u^\star(x,q,W) = \mathcal{L}_x\mathcal{L}_y G(x,q)(\mathcal{L}_x\mathcal{L}_y G(q,q))^{-1} W \tag{32}$$

and the conditional covariance kernel $\mathcal{C}^\omega(x,y) = (16)$ *given by*

$$\mathcal{L}_x\mathcal{L}_y G(x,y) - \mathcal{L}_x\mathcal{L}_y G(x,q)(\mathcal{L}_x\mathcal{L}_y G(q,q))^{-1} \mathcal{L}_x\mathcal{L}_y G(q,y). \tag{33}$$

Proof (30) and (31) follow from the generalized representer theorem (see Corollary 17.12 in Ref. [9] and Proposition 2.11 in Ref. [22]) and the identity $\operatorname{curl}\operatorname{curl}^{\mathrm{T}} = I_3\Delta - \operatorname{Hess}$. $\delta_{k,m}$

represents the Kronecker delta ($\delta_{k,m}=1$ for $k=m$, and $\delta_{k,m}=0$ otherwise). For $\alpha\in(\mathbb{R}^3)^N$, the identity

$$\alpha^{\mathrm{T}}\mathcal{L}_x\mathcal{L}_yG(q,q)\alpha=\sum_{m=1}^{3}\Big\|\sum_{i=1}^{N}\sum_{k=1}^{3}\delta_{q_i}\circ(\alpha_{i,k}(\delta_{k,m}\Delta_x-\partial_{x_k}\partial_{x_m}))\Big\|_G^{*,2}$$

implies that $\mathcal{L}_x\mathcal{L}_yG(q,q)$ is invertible if q_i are pairwise distinct. (30) and the identities $\operatorname{curl}\operatorname{curl}^{\mathrm{T}}=\mathcal{L}$ and $\mathcal{C}^{\omega}(x,y)=\operatorname{curl}_x\operatorname{curl}_y^{\mathrm{T}}\mathcal{C}^{u}(x,y)$ imply (32) and (33).

Using Proposition 4, (14) reduces to

$$\begin{cases}\dot{q}_i^{\star}=\operatorname{curl}_x^{\mathrm{T}}\mathcal{L}_yG(q_i^{\star},q^{\star})(\mathcal{L}_x\mathcal{L}_yG(q^{\star},q^{\star}))^{-1}W^{\star},\\ \dot{W}_i^{\star}=\nu\Delta_x\mathcal{L}_x\mathcal{L}_yG(q_i^{\star},q^{\star})(\mathcal{L}_x\mathcal{L}_yG(q^{\star},q^{\star}))^{-1}W^{\star}\\ \qquad\quad+W_i^{\star}\nabla_x\operatorname{curl}_x^{\mathrm{T}}\mathcal{L}_yG(q_i^{\star},q^{\star})(\mathcal{L}_x\mathcal{L}_yG(q^{\star},q^{\star}))^{-1}W^{\star}+g(q_i^{\star}(t),t).\end{cases}\tag{34}$$

4.3 Periodic kernels

We now describe the construction of the kernel G, which must be a non-generate C^3-differentiable scalar-valued kernel on $\mathbb{T}^d$ such that $\mathcal{H}_G$ is compactly embedded in $H^s(\mathbb{T}^d)$ for $s>4+d/2$. One approach to designing G is to compose a (sufficiently regular and non-degenerate) kernel $\mathfrak{g}$ on $\mathbb{R}^{2d}\times\mathbb{R}^{2d}$ with the function $h:\mathbb{T}^d\to\mathbb{R}^{2d}$ defined by

$$h(x)=(\cos x_1,\sin x_1,\cdots,\cos x_d,\sin x_d),\tag{35}$$

and we obtain

$$G(x,y)=\mathfrak{g}(h(x),h(y)).\tag{36}$$

Taking $\mathfrak{g}$ as the Gaussian kernel, $\mathfrak{g}(X,Y)=\exp\big(-\frac{|X-Y|^2}{2\sigma^2}\big)$ results in

$$G(x,y)=\exp\Big(\frac{-d+\sum_{i=1}^{d}\cos(x_i-y_i)}{\sigma^2}\Big),\tag{37}$$

which satisfies the requirements of G.

Remark 2 Assume $\mathfrak{g}$ to be analytic. It follows that the elements of its RKHS $\mathcal{H}_{\mathfrak{g}}$ are analytic functions[27]. Therefore, for every function of the form $f\circ h$ with $f\in\mathcal{H}_{\mathfrak{g}}$ and $\mathfrak{g}$ being analytic, f is uniquely determined by its values on the range of h. $\langle\cdot,\cdot\rangle_{\mathfrak{g}}$ ($\|\cdot\|_{\mathfrak{g}}$) represents the RKHS inner product (norm) defined by $\mathfrak{g}$. Therefore, for $f\in\mathcal{H}_{\mathfrak{g}}$, we define the norm $\|f\circ h\|:=\|f\|_{\mathfrak{g}}^2$ and use $\langle\cdot,\cdot\rangle$ as its associated inner product. The reproducing property

$$\big\langle f\circ h,\mathfrak{g}(h(\cdot),h(x))\big\rangle=\big\langle f,\mathfrak{g}(\cdot,h(x))\big\rangle_{\mathfrak{g}}=f\circ h(x)\tag{38}$$

for $f\in\mathcal{H}_{\mathfrak{g}}$ implies that $\mathcal{H}_G=\{f\circ h\mid f\in\mathcal{H}_{\mathfrak{g}}\}$ and the RKHS norm $\|\cdot\|_G$ defined by G is $\|\cdot\|$, i.e.,

$$\|f\circ h\|_G^2=\|f\|_{\mathfrak{g}}^2\quad\text{for}\quad f\in\mathcal{H}_{\mathfrak{g}}.\tag{39}$$

If $\mathfrak{g}$ is not analytic, these results are generalized to $\mathcal{H}_G=\{f\circ h\mid f\in\mathcal{H}_{\mathfrak{g}}\}$ with

$$\|v\|_G^2=\inf_{f\in\mathcal{H}_{\mathfrak{g}}:\,f\circ h=v}\|f\|_{\mathfrak{g}}^2.\tag{40}$$

To demonstrate this, we observe that because $\{f\in\mathcal{H}_{\mathfrak{g}}:\,f\circ h=v\}$ is a closed affine subspace of $\mathcal{H}_{\mathfrak{g}}$, the infimum in (40) is achieved and can be expressed as Pv, where P is a linear operator.

Therefore, $\|v\|^2 = \|Pv\|^2_{\mathfrak{g}}$ and $\langle v, v'\rangle = \langle Pv, Pv'\rangle_{\mathfrak{g}}$ define a quadratic norm and an inner product of $\{f \circ h \mid f \in \mathcal{H}_{\mathfrak{g}}\}$, respectively, satisfying the reproducing identity

$$\big\langle v, \mathfrak{g}(h(\cdot), h(x))\big\rangle = \big\langle Pv, P\mathfrak{g}(h(\cdot), h(x))\big\rangle_{\mathfrak{g}} = \big\langle Pv, \mathfrak{g}(\cdot, h(x))\big\rangle_{\mathfrak{g}} = (Pv) \circ h(x) = v(x), \tag{41}$$

which establishes (40). The identity $P\mathfrak{g}(h(\cdot), h(x)) = \mathfrak{g}(\cdot, h(x))$ employed in (41) follows from observing that the identity $\|\mathfrak{g}(\cdot, h(x)) + f\|^2_{\mathfrak{g}} = \mathfrak{g}(h(x), h(x)) + \|f\|^2_{\mathfrak{g}} + 2f \circ h(x)$ implies that the minimizer of $\|\mathfrak{g}(\cdot, h(x)) + f\|^2_{\mathfrak{g}}$ over $f \in \mathcal{H}_{\mathfrak{g}}$ such that $f \circ h = 0$ is $f = 0$.

5 Power-laws informed kernels

We now investigate the incorporation of known scaling and power laws into the selection of the kernel G introduced in Section 4 to derive the divergence-free kernel K. We focus on the two-thirds law derived by Kolmogorov[28] from symmetry and universality assumptions of fully developed (homogeneous and isotropic) turbulence.

When the dimension $d = 3$, the two-thirds law of fully developed (homogeneous and isotropic) turbulence states that the mean of the velocity increment $|u(x + y, t) - u(x, t)|^2$ behaves approximately as $|y|^{\frac{2}{3}}$, which is the two-thirds power of the distance $|y|$ between the points $x + y$ and x (see Chapter 5 in Ref. [29]), which "is equivalent to the statement that the energy spectrum follows a $k^{-\frac{5}{3}}$ law over a suitable range" (see Page 61 in Ref. [29]).

When the dimension $d = 2$, the statistics of the velocity increments follow a different power-law[30–31]: the mean of the squared velocity increment $|u(x + y, t) - u(x, t)|^2$ behaves approximately as $|y|^2$, which is equivalent to the statement that the energy spectrum follows a k^{-3} law (see Page 56 in Ref. [29]).

To incorporate these power laws, we observe that, in the proposed GP approach, the velocity u is randomized according to the distribution of $\xi \sim \mathcal{N}(0, K)$. Therefore, we use the identity

$$\mathbb{E}(|\xi(x) - \xi(y)|^2) = \mathrm{Tr}(K(x, x) + K(y, y) - 2K(x, y)) \tag{42}$$

to incorporate the aforementioned velocity-increment power laws. Considering the scenario in which $G(x, y)$ is stationary (as in (37)), i.e., $G(x, y) = \psi(x - y)$ for a function ψ, (20) and (28) reduce to the particular construction of Subsection 5.1 in Ref. [26], i.e.,

$$K(x, y) = (\mathrm{Hess}\,\psi - \mathrm{Tr}(\mathrm{Hess}\,\psi)I_d)(x - y). \tag{43}$$

Thus, (42) reduces to

$$\mathbb{E}\big(|\xi(x) - \xi(y)|^2\big) = 2(d - 1)(\Delta\psi(x - y) - \Delta\psi(0)). \tag{44}$$

5.1 Richardson cascade

The basic phenomenology of turbulence, known as the Richardson cascade (see Chapter 7 in Ref. [29]), is that the velocity field is composed of space-filling eddies of various sizes $\ell_0 r^n$ for $0 < r < 1$ and $n = 0, 1, \cdots, m$. This phenomenology is associated with the concept of an energy cascade, which indicates that energy is transferred from large (inertial) scales of motion to small (dissipative) scales. 2D turbulence is also associated with the possible presence of an inverse energy cascade[32] representing the transfer of energy from small to large scales. The dissipation scale $\ell_c \sim \ell_0 r^m$ is identified by matching the convective transport time scale $(\ell/\delta v(\ell))$ to the diffusive transport time scale (ℓ^2/ν). For $d = 3$ (using $\delta v(\ell) \sim U_0(\ell/\ell_0)^{\frac{1}{3}}$), this translates to $\ell_c/\ell_0 \sim Re^{-\frac{3}{4}}$, where $Re = U_0\ell_0/\nu$ is the Reynolds number. For $d = 2$ (using $\delta v(\ell) \sim U_0(\ell/\ell_0)$), this translates to $\ell_c/\ell_0 \sim Re^{-\frac{1}{2}}$. We incorporate these concepts from a statistical perspective by representing the GP $\xi \sim \mathcal{N}(0, K)$ as an additive GP

$$\xi = \sum_{n=0}^{m} \xi^{(n)}, \tag{45}$$

where $\xi^{(n)} \sim \mathcal{N}(0, K^{(n)})$ are independent and represent eddies on a scale indexed by n. Representing ξ as an additive GP is equivalent to representing K as an additive kernel

$$K = \sum_{n=0}^{m} K^{(n)}. \tag{46}$$

To ensure that K and $K^{(n)}$ are divergence-free matrix valued kernels, we select, as in Section 4, $K^{(n)}(x,y) := \operatorname{curl}_x^{\mathrm{T}} \operatorname{curl}_y G^{(n)}(x,y)$, where $G^{(n)}$ is a periodic scalar-valued kernel on $\mathbb{T}^d$. This is equivalent to selecting $K(x,y) := \operatorname{curl}_x^{\mathrm{T}} \operatorname{curl}_y G(x,y)$ with

$$G = \sum_{n=0}^{m} G^{(n)}. \tag{47}$$

5.2 Power laws

We now incorporate the velocity-increment power laws into the selection of kernels $G^{(n)}$. To incorporate periodicity, stationarity, power laws, and self-similarity, we select (as in (37))

$$G^{(n)}(x,y) = \alpha_n \exp\left(\frac{-d + \sum_{i=1}^{d} \cos(x_i - y_i)}{\sigma_n^2}\right) \tag{48}$$

with

$$\sigma_n = \frac{\sigma_0}{2^n}, \quad \alpha_n = \sigma_n^{\gamma} \tag{49}$$

for $\gamma \in \mathbb{R}$ to be determined using the power-law discussed in Section 5. For

$$\psi^{(n)}(x) := \alpha_n \exp\left(\frac{-d + \sum_{i=1}^{d} \cos x_i}{\sigma_n^2}\right), \tag{50}$$

we obtain

$$\Delta\psi^{(n)}(x) := \sum_{j=1}^{d} \alpha_n \Big(\frac{\sin^2 x_j}{\sigma_n^4} - \frac{\cos x_j}{\sigma_n^2}\Big) \exp\left(\frac{-d + \sum_{i=1}^{d} \cos x_i}{\sigma_n^2}\right), \tag{51}$$

which, by using (44), leads to

$$\mathbb{E}(|\xi(x) - \xi(0)|^2) = 2(d-1) \sum_{n=0}^{m} \alpha_n \sum_{j=1}^{d} \left(\frac{1}{\sigma_n^2} + \Big(\frac{\sin^2 x_j}{\sigma_n^4} - \frac{\cos x_j}{\sigma_n^2}\Big) \exp\left(\frac{-d + \sum_{i=1}^{d} \cos x_i}{\sigma_n^2}\right)\right). \tag{52}$$

We deduce that for $x \sim 2^{-q}$ with $1 < q < m$,

$$\mathbb{E}(|\xi(x) - \xi(0)|^2) \sim 2(d-1) \sum_{n=q}^{m} \alpha_n \frac{d}{\sigma_n^2}. \tag{53}$$

Observing that $\sigma_n \sim 2^{-n}$ and $\alpha_n \sim 2^{-n\gamma}$, it follows that for $|x| \sim 2^{-q}$ and $\gamma > 2$,

$$\mathbb{E}(|\xi(x) - \xi(0)|^2) \sim 2^{q(2-\gamma)}. \tag{54}$$

Therefore, the velocity-increment power laws in Section 5 are incorporated by obtaining

$$\begin{cases} \gamma = 4 \quad \text{for} \quad d = 2, \\ \gamma = \dfrac{2}{3} + 2 \quad \text{for} \quad d = 3. \end{cases} \tag{55}$$

5.3 Mode decomposition

Although the Richardson cascade is based on a qualitative analysis of turbulence supported by the qualitative notion of eddies at different scales, this analysis can be performed quantitatively by using kernel mode decomposition[11]. To describe this, we observe that the decomposition (45) leads to a corresponding decomposition of the velocity field (11), i.e.,

$$u^\star(x,q,W) = \sum_{n=0}^{m} u^{(n)}(x,q,W) \tag{56}$$

with

$$u^{(n)}(x,q,W) = \mathbb{E}(\xi^{(n)}(x)\big|\operatorname{curl}\xi(q) = W), \tag{57}$$

where $u^{(n)}$ are the following representer formulae (using the notations of (22) and (30)):

$$u^{(n)}(x,q,W) = \begin{cases} \operatorname{curl}_x^{\mathrm{T}} \Delta_y G^{(n)}(x,q)(\Delta_x\Delta_y G(q,q))^{-1}W, & d=2,\\ \operatorname{curl}_x^{\mathrm{T}} \mathcal{L}_y G^{(n)}(x,q)(\mathcal{L}_x\mathcal{L}_y G(q,q))^{-1}W, & d=3. \end{cases} \tag{58}$$

Furthermore, the RKHS norm of $u^\star$ admits the decomposition[11]

$$\|u^\star\|_K^2 = \sum_{n=0}^{m} \|u^{(n)}\|_{K^{(n)}}^2, \tag{59}$$

where

$$\|u^{(n)}\|_{K^{(n)}}^2 = \left\langle u^{(n)}, u^*\right\rangle_K = \operatorname{Var}\left[\left\langle \xi^{(n)}, u^*\right\rangle_K\right] \tag{60}$$

can be interpreted as a measure of the activation of the GP (mode) $\xi^{(n)}$ after conditioning on $\xi(q) = W$. Using $K^{(n)}(x,y) := \operatorname{curl}_x^{\mathrm{T}} \operatorname{curl}_y G^{(n)}(x,y)$, we obtain

$$\left\|u^{(n)}(\cdot,q,W)\right\|_{K^{(n)}}^2 = \begin{cases} W^{\mathrm{T}}(\Delta_x\Delta_y G(q,q))^{-1}W, & d=2,\\ W^{\mathrm{T}}(\mathcal{L}_x\mathcal{L}_y G(q,q))^{-1}W, & d=3. \end{cases} \tag{61}$$

6 Accuracy of the proposed approach and information loss

6.1 Residual (source term error) as a measure of accuracy

The accuracy of the proposed approach can be characterized by two terms. The first one is the error $\overline{\omega}(x,0) - \omega(x,0)$ in approximating the initial value of the vorticity. The second term is the spurious source term $\mathfrak{s}$ introduced by the numerical method, defined as (see Fig. 7 for snapshots, where we use periodic boundary conditions such that the errors in these snapshots are solely a reflection of the particle locations and the initial condition) $|\mathfrak{s}(\cdot,t)|$)

$$\mathfrak{s}(x,t) := \partial_t\overline{\omega} + \overline{u}\nabla\overline{\omega} - \nu\Delta\overline{\omega} - g(x,t), \quad d=2, \tag{62}$$

$$\mathfrak{s}(x,t) := \partial_t\overline{\omega} + \overline{u}\nabla\overline{\omega} - \nu\Delta\overline{\omega} - \overline{\omega}\nabla\overline{u} - g(x,t), \quad d=3. \tag{63}$$

The first term $\overline{\omega}(x,0) - \omega(x,0)$ is well-understood as a kernel interpolation error, and a-priori error estimates can be obtained from Poincaré inequalities[9,33]: the norm of this term can be shown to decay towards zero as a power of the fill distance between the collocation points $q_i(0)$ (the power depends on the strength of the norm, the regularity of ω_0, and the regularity of the RKHS defined by the GP curl $\xi(x)$; see Refs. [9] and [33] for details and further references).

The second term $\mathfrak{s}(x,t)$ is not well-understood and we seek to analyze it. Note that this term is zero at the particle locations $q_i(t)$ ($\mathfrak{s}(q_i(t),t) = 0$) and is a function of the choice of the

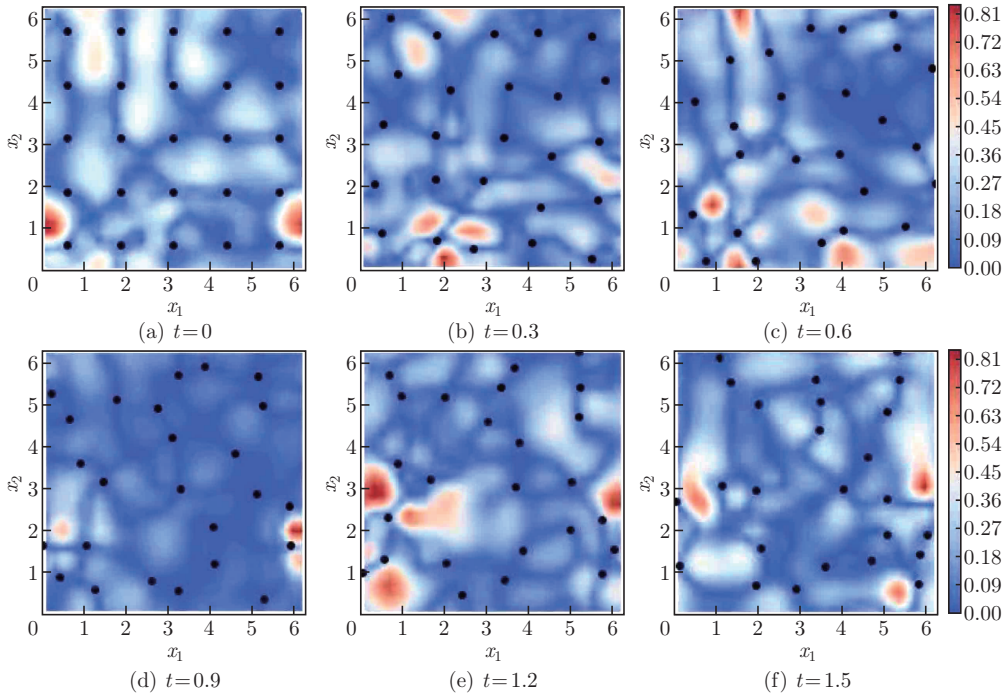

Fig. 7 Source terms error $\mathfrak{s}(\cdot,t)$ (color online)

kernel for ξ and the number of particles N. Although the stability estimates (they are available for $d=2$[34] but remain a challenge for $d=3$[35]) for NS equations enable us to determine the norm of the errors on velocity $u-\overline{u}$ and vorticity $\omega-\overline{\omega}$, we do not expect these bounds to be useful because the chaotic nature of the NS equations would imply their rapid blow-up as a function of time (caused by a blow-up of the stability constants) in turbulent regimes. However, $\mathfrak{s}(x,t)$ is a more useful measure of error since it acts as an instantaneous error made on the source term of the NS equations by the proposed numerical method: modulo the initial value error $\omega(x,0)-\overline{\omega}(x,0)$, simulating $\overline{\omega}$ is equivalent to simulating the continuous NS equations with the added source term $\mathfrak{s}(x,t)$.

6.2 $\mathfrak{s}$ as a measure of information loss

$\mathfrak{s}$ can also be interpreted as a measure of the information loss. To describe this, let $t_0\geqslant 0$ and $q_a(t)$ be the trajectory of the particle driven by the flow $\overline{u}(x,t)$ ($\dot{q}_a(t)=\overline{u}(q_a(t),t)$) and starting at time t_0 at an arbitrary point $x\in\mathbb{T}^d$. Let $W_a(t)=\overline{\omega}(q_a(t),t)$ be the predicted vorticity at $q_a(t)$. Let $q_e:=(q^\star,q_a)$ (respectively, $W_e:=(W^\star,W_a)$) be the vector of particle locations obtained by concatenating $q^\star$ with q_a (respectively, $W^\star$ with W_a). Thus, the identity

$$u^\star(x,q^\star,W^\star)=u^\star(x,q_e,W_e) \tag{64}$$

implies that (q_a,W_a) does not carry (additional) information on the approximation of the flow, given the information contained in $(q^\star,W^\star)$. Now, let W_b be the solution to

$$\dot{W}_b(t)=\nu\Delta\operatorname{curl}u^\star(q_a(t),q^\star,W^\star)+g(q_a(t),t),\quad d=2, \tag{65}$$

$$\dot{W}_b(t)=\nu\Delta\operatorname{curl}u^\star(q_e(t),q^\star,W^\star)+W_b(t)\nabla u^\star(q_e(t),q^\star,W^\star)+g(q_e(t),t),\quad d=3 \tag{66}$$

with the initial condition $W_b(t_0)=W_a(t_0)$. Thus, the identity

$$W_b(t)-W_a(t)=\mathfrak{s}(x,t)(t-t_0)+o(t-t_0) \tag{67}$$

implies that $|\mathfrak{s}(x,t)|$ can be interpreted as the instantaneous rate of information gain at time t_0 resulting from adding a particle at x and letting W_b be driven by the GPH equations. Equivalently, $|\mathfrak{s}(x,t)|$ can be interpreted as the rate of information loss resulting from the absence of an additional particle at location x. Therefore, to minimize information loss, the number of particles in GPH can be dynamically increased by adding new particles at the location x, where $|\mathfrak{s}(x,t)|$ is maximized (a similar concept of information loss can be derived for particle removal).

6.3 Turbulence as information loss

How do we define and quantify turbulence? The current popular definition, "the complex, chaotic motion of a fluid"[2], is not only empirical but also relative to the scale at which the flow is observed (the flow may appear laminar at fine scales and chaotic at coarse scales). From the GPH perspective, turbulence can be defined as the information loss incurred by approximating the dynamics of a continuous flow with the discrete information contained in $(q^\star, W^\star)$. In this sense, it is a local quantity measured as $|\mathfrak{s}(x,t)|$ and its definition is relative to the information already contained $(q^\star, W^\star)$.

7 Numerical experiments

In the following experiments, for the dimension $d = 2$, we use the additive kernel of Subsections 5.1 and 5.2 with $m + 1 = 3$ modes, $\gamma = 4$, and $(\sigma_0, \sigma_1, \sigma_2) = (2, 1, 0.5)$. We use $N = 25$ particles, zero-forcing ($f = 0$), and zero viscosity ($\nu = 0$), and we initialize the vorticity field at random by sampling the initial value of W from the distribution of the Gaussian vector with the identity covariance matrix. Figures 1 and 2 show the snapshots of the velocity field $(x \to \overline{u}(x,t))$ and the vorticity field $(x \to \overline{\omega}(x,t))$ with the entries of $q(t)$ shown as particles. Figures 4 and 5 show the snapshots of the variance of the velocity field $(x \to \mathrm{Tr}[\mathcal{C}^u(x,x)])$ and the variance of the vorticity field $(x \to \mathrm{Tr}[\mathcal{C}^\omega(x,x)])$. Figure 7 shows the snapshots of the source term error $(x \to \mathfrak{s}(x,t))$. Figure 8(a) shows the mode activation of each of the three modes as defined in (61). Figure 8(b) shows the power spectrum of the field generated by our simulation and its comparison with the k^{-3} power spectrum associated with 2D turbulence, where k is the frequency. Figure 8(c) shows the source term error $t \to \|\mathfrak{s}(\cdot,t)\|_{L^2}$, where $\|\mathfrak{s}(\cdot,t)\|_{L^2}^2 := |\mathbb{T}^d|^{-1} \int_{\mathbb{T}^d} \mathfrak{s}^2(x,t)\,\mathrm{d}x$. The plots shown in Fig. 8 are for zero viscosity $\nu = 0$. Figures 9 and 10 present similar plots for $\nu = 0.001$ and $\nu = 0.01$, respectively. Figure 11 shows similar plots for $\nu = 0$, $m + 1 = 5$ modes, $\gamma = 4$, and $(\sigma_0, \sigma_1, \sigma_2, \sigma_3, \sigma_4) = (2, 1, 1/2, 1/4, 1/8)$. Note that compared with Fig. 8(c), the source term error $t \to \|\mathfrak{s}(\cdot,t)\|_{L^2}$ is decreased by one order of magnitude, which supports the point that our structured multiscale kernel leads to increased accuracy as the number of modes is increased.

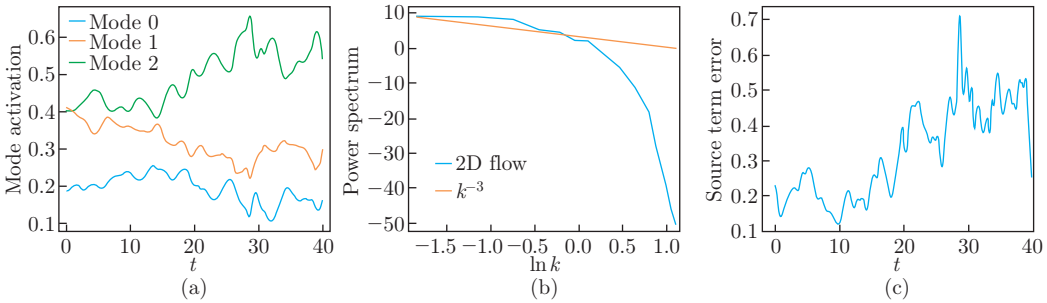

Fig. 8 (a) Mode activation, (b) power spectrum, and (c) source term error $t \to \|\mathfrak{s}(\cdot,t)\|_{L^2}$ when $\nu = 0$ (color online)

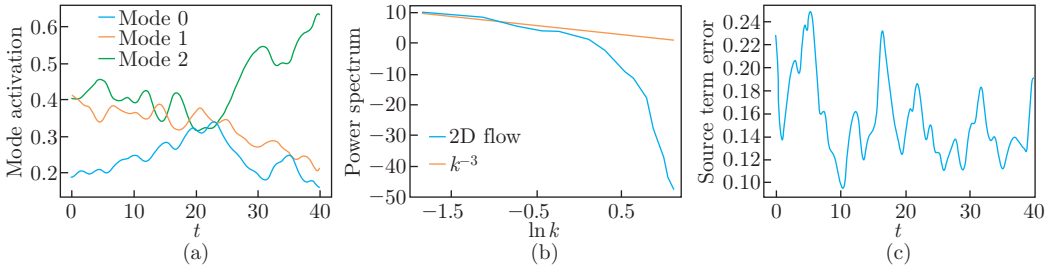

Fig. 9 (a) Mode activation, (b) power spectrum, and (c) source term error $t \to \|\mathfrak{s}(\cdot,t)\|_{L^2}$ when $\nu = 0.001$ (color online)

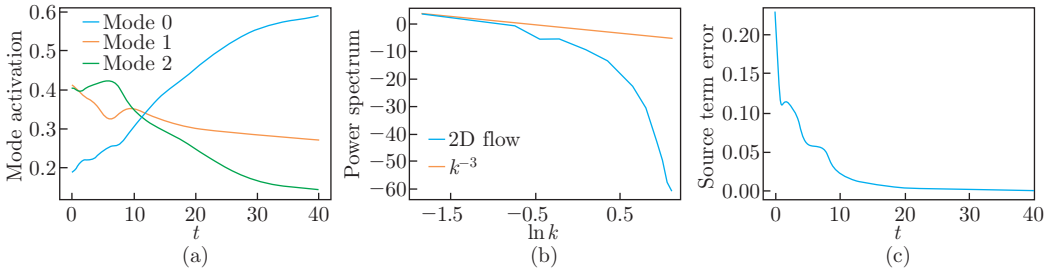

Fig. 10 (a) Mode activation, (b) power spectrum, and (c) source term error $t \to \|\mathfrak{s}(\cdot,t)\|_{L^2}$ when $\nu = 0.01$ (color online)

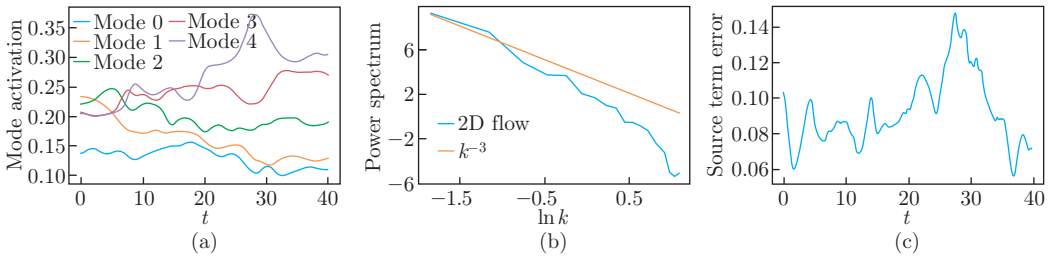

Fig. 11 (a) Mode activation with 5 modes, (b) power spectrum, and (c) source term error $t \to \|\mathfrak{s}(\cdot,t)\|_{L^2}$ when $\nu = 0$ (color online)

Table 1 provides the space/time-averaged source term error ($T = 40$)

$$\|\mathfrak{s}\|_{L^2} := \sqrt{T^{-1} \int_{\mathbb{T}^d \times [0,T]} \mathfrak{s}^2(x,t)\, \mathrm{d}x\, \mathrm{d}t} \tag{68}$$

as a function of the number of modes ($m+1$ in the additive kernel of Subsections 5.1 and 5.2)

Table 1 Space/time average source term error $\|\mathfrak{s}\|_{L^2}$ as a function of the number of modes $m+1$ and the power-law parameter γ

$m+1$	γ				
	-2	0	2	4	6
2	27.6	39.4	93.4	31.1	34.5
3	0.118	0.158	0.214	0.389	1.144
4	1.14	4.22	0.076	0.070	0.266

and the value of the parameter γ in the power law (49). For Table 1, we use $N = 100$ particles, zero-forcing ($f = 0$), and $\nu = 0.001$, and initialize the vorticity field at random by sampling the initial value of W from the distribution of the Gaussian vector with the identity covariance matrix. Selecting the parameter γ close to the one ($\gamma = 4$) matching the Kolmogorov scaling law and increasing the number of modes $m + 1$ significantly diminishes the source term error $\|\mathfrak{s}\|_{L^2}$. With only one mode ($m + 1 = 1$, not shown in Table 1), the kernel K is too stiff to handle the transfer of energy towards fine scales, and the accuracy significantly deteriorates ($\|\mathfrak{s}\|_{L^2} \sim 17\,000$, and without regularization with a nugget, the velocity bursts are observed as particles come close to each other).

Remark 3 The values of $\|\mathfrak{s}\|_{L^2}$ are absolute in Table 1 and Figs. 7–11. Our main purpose is to show the dependence of $\|\mathfrak{s}\|_{L^2}$ as a function of the number of modes and the value of the parameter γ. In particular, those values can be made relative by dividing them by 17 000 (the value of $\|\mathfrak{s}\|_{L^2}$ with only one mode).

Figure 12 illustrates the convergence of the method (as measured by $\|\mathfrak{s}\|_{L^2}$) with respect to the number of particles N. In the figure, when $W(0) = \omega_0(q(0))$, the initial vorticity ω_0 is selected to be smooth and deterministic. The interpolation error in the approximation of the initial vorticity is not plotted (the analysis of this kernel interpolation error is classical[9]).

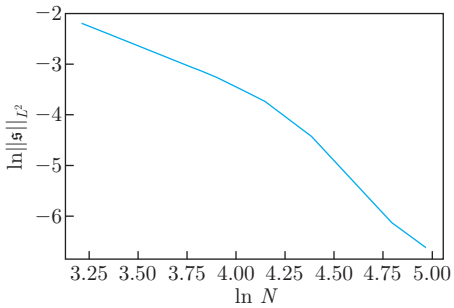

Fig. 12 Error $\|\mathfrak{s}\|_{L^2}$ vs. number of particles N in the log-log scale (color online)

Remark 4 The complexity of the method is proportional to the product of the number of time steps and the cost of inverting dense $N \times N$ kernel matrices. Although the sparse Cholesky factorization algorithms introduced in Refs. [19] and [20] can be adapted to potentially reduce the inversion cost to $\mathcal{O}(N \ln^{2d} N)$, we have not employed this strategy here.

For the 3D setting ($d = 3$), we also use the additive kernel of Subsections 5.1 and 5.2 with $m + 1 = 1$ and $m + 1 = 2$ modes, $\gamma = 2/3 + 2$, and $(\sigma_0, \sigma_1) = (2, 1)$. We use $N = 9$ particles, zero-forcing ($f = 0$), and non-zero viscosity ($\nu = 0.001$), and initialize the vorticity field at random by sampling the initial value of W from the distribution of the Gaussian vector with the identity covariance matrix. Figures 3 and 13 show the snapshots of the velocity field ($x \to \overline{u}(x, t)$) and the vorticity W at locations $q(t)$. Figures 3 and 13 employ one and two modes, respectively. The added mode increases the effective viscosity of the dynamics by acting as an energy sink. Compared with the 2D setting, the 3D ODE formulation of GPH has a quadratic term in $W^\star$ in (34) that can lead to a blowup in finite time. We numerically observe this blowup and dampen the vortex stretching component of this quadratic term (using $w_i = w_{i,\parallel} + w_{i,\perp}$ for the orthogonal decomposition of $W_i^\star \nabla_x \operatorname{curl}_x^{\mathrm{T}} \mathcal{L}_y G(q_i^\star, q^\star)\, (\mathcal{L}_x \mathcal{L}_y G(q^\star, q^\star))^{-1} W^\star$ into its projection along the direction of $W_i^\star$ and its orthogonal complement, we replace w_i by $w_i = (1-\alpha) w_{i,\parallel} + w_{i,\perp}$) by a factor $1-\alpha$ (with $\alpha \in [0, 1)$) to avoid blowup. Other strategies for avoiding blowup in the numerical calculation of the NS and Euler equations include numerical dissipation and Lagrangian averaging[36]. Although the 3D Euler equations with boundary and smooth initial data can blow up, the blowup of the 3D NS equations remains an open problem.

Therefore, addressing the possible blowup of (34) in a manner that has better consistency with the underlying physics of turbulence remains an open problem that may require a modeling step (i.e., correcting the NS equations). We also note that if the continuous 3D NS equations do indeed blow up, the solution obtained with GPH will exit the RKHS defined by a smooth kernel. Finally, GPH may also serve as a potential candidate for identifying a singularity formation in the solution to the 3D NS equations: if a trajectory $(q^\star, W^\star)$ and a (possibly time-dependent) kernel G exist, such that $W^\star$ blows up in finite time, whereas $\mathfrak{s}$ in (63) (with $g = 0$) remains smooth. Then, the NS equations blow up in finite time[1].

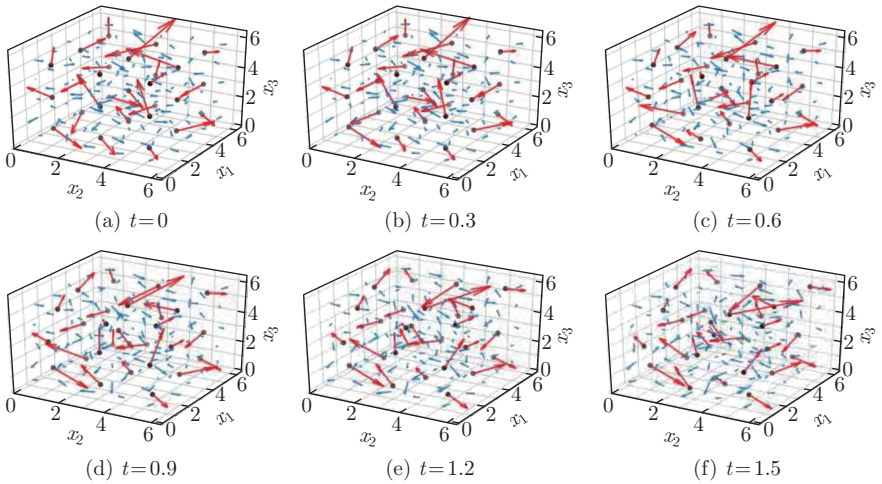

Fig. 13 Velocity and vorticity snapshots for $d = 3$, where the blue arrows indicate the velocity, the red arrows indicate the vorticity W at particle locations q, and the kernel has 2 modes (color online)

8 Further discussion

8.1 Selecting the kernel when the physics is unknown

The approach proposed in this paper involves designing the kernel to satisfy known physics. When the underlying physics is unknown, the kernel can be learned from data via cross-validation/maximum likelihood estimation in a given (possibly non-parametric) family of kernels[25,37–38]. The kernel flow (a variant of cross-validation) approach[37] has been shown to be efficient for learning (possibly stochastic) dynamical systems[39–43] and designing surrogate models[44–46]. In particular, this approach has been shown to compare favorably to ANN-based methods (in terms of both complexity and accuracy) for weather/climate prediction using actual satellite data[40].

8.2 GPH and ANN-based simulations

The aim of this manuscript is not to compare GPH against ANN-based methods for solving the NS equations (refer to Ref. [22] for such comparisons for general PDEs) but to highlight that GP-based methods enable the incorporation of physics, and not solely by enforcing the PDE for a finite number of collocation points/particles but also through the choice and design of the kernel. Thus, our analysis and results can be extended to derive an ANN variant of GPH. This variant can be obtained by simply defining the scalar-valued kernel G introduced in Section 4 as

$$G(x, x') = \psi_\theta^{\mathrm{T}}(x)\psi_\theta(x'), \tag{69}$$

where $\psi_\theta(x)$ is the output of an ANN, i.e., a function mapping x to a finite-dimensional vector space parameterized by the parameters θ in the inner layers of a neural network. (69) then defines a parameterized kernel whose parameters can be learned from data, as described in Subsection 8.1.

8.3 UQ and data assimilation

Because the proposed approach uses a Bayesian interpretation, it naturally enables data assimilation and predictions and estimations by mixing simulation and experimental data. To describe this, we assume that, in addition to the information (q, W) obtained from the simulation, we have access (as functions of time) to velocity measurements $v_1, v_2, \cdots, v_M$ at locations $z_1, z_2, \cdots, z_M$ (that may be time-dependent). Thus, GPH can be modified to incorporate this information. To describe this, we write

$$u^\star(x, q, W, z, v) := \mathbb{E}(\xi(x)\big|\operatorname{curl}\xi(q) = W \text{ and } \xi(z) = v). \tag{70}$$

This modification can then be summarized as approximating $u(x,t)$ with

$$\overline{u}(x,t) := u^\star(x, q^\star(t), W^\star(t), z, v), \tag{71}$$

and $\omega(x,t)$ with

$$\overline{\omega}(x,t) := \operatorname{curl} u^\star(x, q^\star(t), W^\star(t), z, v), \tag{72}$$

where $(q^\star, W^\star)$ is the solution to the autonomous system of ODEs

$$\begin{cases} \dot{q}_i^\star = u^\star(q_i^\star, q^\star, W^\star, z, v), \\ \dot{W}_i^\star(t) = \nu\Delta\operatorname{curl} u^\star(q_i^\star, q^\star, W^\star, z, v) + g(q_i^\star(t), t), \quad d = 2, \\ \dot{W}_i^\star(t) = \nu\Delta\operatorname{curl} u^\star(q_i^\star, q^\star, W^\star, z, v) + W_i^\star\nabla u^\star(q_i^\star, q^\star, W^\star, z, v) + g(q_i^\star(t), t), \quad d = 3 \end{cases} \tag{73}$$

with the initial condition $(q^\star, W^\star)(0) = (q, W)(0) = (q^0, \omega_0(q^0))$. Note that this modification is equivalent to replacing the distribution of the GP ξ in Section 3 by that of a non-centered time dependent GP with the mean $\mathbb{E}(\xi(x)\big|\xi(z) = v)$ and the covariance function defined as the conditional covariance of ξ conditioned on $\xi(z) = v$. Representer formulae can be obtained naturally as described in Section 4. Other experimental measurements may also be incorporated (e.g., vorticities at specific locations). Furthermore, using the proposed approach, the velocity and pressure fields can be learned from flow visualizations as in Ref. [47], with the advantage of recovering uncertainties (whole posterior distributions) in addition to those fields.

To describe this, we assume that we have access (as functions of time) to the values $y_1, y_2, \cdots, y_M$ at locations $z_1, z_2, \cdots, z_M$ of the concentration c of a passive tracer satisfying the transport PDE,

$$\partial_t c + u \cdot \nabla c = D\Delta c. \tag{74}$$

Let Γ be a smoothing scalar valued kernel and $\zeta \sim \mathcal{N}(0, \Gamma)$. Write

$$\overline{c}(x,t) := \mathbb{E}(\zeta(x)|\zeta(z) = y(t)), \tag{75}$$

and

$$u^\star(x, q, W, t) := \mathbb{E}(\xi(x)\big|\operatorname{curl}\xi(q) = W, \quad \partial_t\overline{c}(z,t) + \xi(z)\cdot\nabla\overline{c}(z,t) = D\Delta\overline{c}(z,t)). \tag{76}$$

$u(x,t)$ can then be approximated with

$$\overline{u}(x,t) := u^\star(x, q^\star(t), W^\star(t), t), \tag{77}$$

and $\omega(x,t)$ can be approximated with

$$\overline{\omega}(x,t) := \operatorname{curl} u^\star(x, q^\star(t), W^\star(t), t), \tag{78}$$

where $(q^\star, W^\star)$ is the solution to the autonomous system of ODEs

$$\begin{cases} \dot{q}_i^\star = u^\star(q_i^\star, q^\star, W^\star, t), \\ \dot{W}_i^\star(t) = \nu\Delta \operatorname{curl} u^\star(q_i^\star, q^\star, W^\star, t) + g(q_i^\star(t), t), \quad d = 2, \\ \dot{W}_i^\star(t) = \nu\Delta \operatorname{curl} u^\star(q_i^\star, q^\star, W^\star, t) + W_i^\star \nabla u^\star(q_i^\star, q^\star, W^\star, t) + g(q_i^\star(t), t), \quad d = 3 \end{cases} \tag{79}$$

with the initial condition $q^\star(0) = q^0$ and

$$W^\star(0) = \mathbb{E}(\operatorname{curl} \xi(q^0) \big| \partial_t \overline{c}(z,0) + \xi(z) \cdot \nabla \overline{c}(z,0) = D\Delta \overline{c}(z,0)). \tag{80}$$

Conflict of interest The author declares no conflict of interest.

Acknowledgements This project was supported by the Air Force Office of Scientific Research under the MURI award number FA9550-20-1-0358 (Machine Learning and Physics-Based Modeling and Simulation) and by the Department of Energy under the award number DE-SC0023163 (SEA-CROGS: Scalable, Efficient, and Accelerated Causal Reasoning Operators, Graphs and Spikes for Earth and Embedded Systems). The author would also like to thank two anonymous referees for their comments and suggestions.

References

[1] FEFFERMAN, C. L. Existence and smoothness of the Navier-Stokes equation. *The Millennium Prize Problems*, **57**, 67 (2000)

[2] PHILLIPS, L. Turbulence, the oldest unsolved problem in physics. Retrieved from arstechnica.com/science/2018/10/turbulence-the-oldest-unsolved-problem-in-physics. Accessed October, **25**, 2018 (2018)

[3] LUCY, L. B. A numerical approach to the testing of the fission hypothesis. *The Astronomical Journal*, **82**, 1013–1024 (1977)

[4] GINGOLD, R. A. and MONAGHAN, J. J. Smoothed particle hydrodynamics: theory and application to non-spherical stars. *Monthly Notices of the Royal Astronomical Society*, **181**, 375–389 (1977)

[5] MONAGHAN, J. J. Smoothed particle hydrodynamics. *Annual Review of Astronomy and Astrophysics*, **30**, 543–574 (1992)

[6] LIU, M. and LIU, G. Smoothed particle hydrodynamics (SPH): an overview and recent developments. *Archives of Computational Methods in Engineering*, **17**, 25–76 (2010)

[7] VACONDIO, R., ALTOMARE, C., DE LEFFE, M., HU, X., LE TOUZÉ, D., LIND, S., MARONGIU, J. C., MARRONE, S., ROGERS, B. D., and SOUTO-IGLESIAS, A. Grand challenges for smoothed particle hydrodynamics numerical schemes. *Computational Particle Mechanics*, **8**, 575–588 (2021)

[8] OWHADI, H., SCOVEL, C., and SCHÄFER, F. Statistical numerical approximation. *Notices of the American Mathematical Society*, **66**, 1608–1617 (2019)

[9] OWHADI, H. and SCOVEL, C. *Operator-Adapted Wavelets, Fast Solvers, and Numerical Homogenization: From a Game Theoretic Approach to Numerical Approximation and Algorithm Design*, Vol. 35, Cambridge University Press, Cambridge (2019)

[10] MICCHELLI, C. A. and RIVLIN, T. J. A survey of optimal recovery. *Optimal Estimation in Approximation Theory*, Springer, New York, 1–54 (1977)

[11] OWHADI, H., SCOVEL, C., and YOO, G. R. *Kernel Mode Decomposition and the Programming of Kernels*, Springer, Cham (2021)

[12] LEONARD, A. Vortex methods for flow simulation. *Journal of Computational Physics*, **37**, 289–335 (1980)

[13] COTTET, G. H. and KOUMOUTSAKOS, P. D. *Vortex Methods: Theory and Practice*, Vol. 8, Cambridge University Press, Cambridge (2000)

[14] RAISSI, M., PERDIKARIS, P., and KARNIADAKIS, G. E. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, **378**, 686–707 (2019)

[15] KARNIADAKIS, G. E., KEVREKIDIS, I. G., LU, L., PERDIKARIS, P., WANG, S., and YANG, L. Physics-informed machine learning. *Nature Reviews Physics*, **3**, 422–440 (2021)

[16] OWHADI, H. Bayesian numerical homogenization. *Multiscale Modeling & Simulation*, **13**, 812–828 (2015)

[17] OWHADI, H. and ZHANG, L. Gamblets for opening the complexity-bottleneck of implicit schemes for hyperbolic and parabolic ODEs/PDEs with rough coefficients. *Journal of Computational Physics*, **347**, 99–128 (2017)

[18] OWHADI, H. Multigrid with rough coefficients and multiresolution operator decomposition from hierarchical information games. *SIAM Review*, **59**, 99–149 (2017)

[19] SCHÄFER, F., KATZFUSS, M., and OWHADI, H. Sparse cholesky factorization by Kullback-Leibler minimization. *SIAM Journal on Scientific Computing*, **43**, A2019–A2046 (2021)

[20] SCHÄFER, F., SULLIVAN, T. J., and OWHADI, H. Compression, inversion, and approximate pca of dense kernel matrices at near-linear computational complexity. *Multiscale Modeling & Simulation*, **19**, 688–730 (2021)

[21] YOO, G. R. and OWHADI, H. De-noising by thresholding operator adapted wavelets. *Statistics and Computing*, **29**, 1185–1201 (2019)

[22] CHEN, Y., HOSSEINI, B., OWHADI, H., and STUART, A. M. Solving and learning nonlinear PDEs with Gaussian processes. *Journal of Computational Physics*, **447**, 110668 (2021)

[23] OWHADI, H. Computational graph completion. *Research in the Mathematical Sciences*, **9**, 1–33 (2022)

[24] BABUŠKA, I. and OSBORN, J. E. Can a finite element method perform arbitrarily badly? *Mathematics of Computation*, **69**, 443–462 (2000)

[25] OWHADI, H. Do ideas have shape? Idea registration as the continuous limit of artificial neural networks. *arXiv Preprint*, arXiv:2008.03920 (2020) https://doi.org/10.48550/arXiv.2008.03920

[26] ALVAREZ, M. A., ROSASCO, L., and LAWRENCE, N. D. Kernels for vector-valued functions: a review. *Foundations and Trends in Machine Learning*, **4**, 195–266 (2011)

[27] SUN, H. W. and ZHOU, D. X. Reproducing kernel hilbert spaces associated with analytic translation-invariant mercer kernels. *Journal of Fourier Analysis and Applications*, **14**, 89–101 (2008)

[28] KOLMOGOROV, A. N. The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers. *Doklady Akademii Nauk SSSR*, **30**, 301–305 (1941)

[29] FRISCH, U. Turbulence: the legacy of A. N. Kolmogorov. *Physics Today*, **49**, 82 (1996)

[30] LINDBORG, E. Can the atmospheric kinetic energy spectrum be explained by two-dimensional turbulence? *Journal of Fluid Mechanics*, **388**, 259–288 (1999)

[31] BOFFETTA, G. and ECKE, R. E. Two-dimensional turbulence. *Annual Review of Fluid Mechanics*, **44**, 427–451 (2012)

[32] SOMMERIA, J. Experimental study of the two-dimensional inverse energy cascade in a square box. *Journal of Fluid Mechanics*, **170**, 139–168 (1986)

[33] WENDLAND, H. Scattered data approximation. *Cambridge Monographs on Applied and Computational Mathematics*, Vol. 17, Cambridge University Press, Cambridge (2005)

[34] ZADRZYNSKA, E. and ZAJCZKOWSKI, W. M. Stability of two-dimensional Navier-Stokes motions in the periodic case. *Journal of Mathematical Analysis and Applications*, **423**, 956–974 (2015)

[35] LADYZHENSKAYA, O. A. Sixth problem of the millennium: Navier-Stokes equations, existence and smoothness. *Russian Mathematical Surveys*, **58**, 251 (2003)

[36] MARSDEN, J. E. and SHKOLLER, S. The anisotropic Lagrangian averaged Euler and Navier-Stokes equations. *Archive for Rational Mechanics and Analysis*, **166**, 27–46 (2003)

[37] OWHADI, H. and YOO, G. R. Kernel flows: from learning kernels from data into the ABYSS. *Journal of Computational Physics*, **389**, 22–47 (2019)

[38] CHEN, Y., OWHADI, H., and STUART, A. Consistency of empirical bayes and kernel flow for hierarchical parameter estimation. *Mathematics of Computation*, **90**, 2527–2578 (2021)

[39] HAMZI, B. and OWHADI, H. Learning dynamical systems from data: a simple cross-validation perspective, part i: parametric kernel flows. *Physica D*: *Nonlinear Phenomena*, **421**, 132817 (2021)

[40] HAMZI, B., MAULIK, R., and OWHADI, H. Simple, low-cost and accurate data-driven geophysical forecasting with learned kernels. *Proceedings of the Royal Society A*, **477**, 20210326 (2021)

[41] DARCY, M., HAMZI, B., SUSILUOTO, J., BRAVERMAN, A., and OWHADI, H. Learning dynamical systems from data: a simple cross-validation perspective, part ii: nonparametric kernel flows. preprint (2021) https://doi.org/10.13140/RG.2.2.16391.32164

[42] LEE, J., DE BROUWER, E., HAMZI, B., and OWHADI, H. Learning dynamical systems from data: a simple cross-validation perspective, part iii: irregularly-sampled time series. *arXiv Preprint*, arXiv:2111.13037 (2021) https://doi.org/10.48550/arXiv.2111.13037

[43] DARCY, M. D., HAMZI, B., LIVIERI, G., OWHADI, H., and TAVALLALI, P. One-shot learning of stochastic differential equations with data adapted kernels. *Physica D*: *Nonlinear Phenomena*, **444**, 133583 (2023)

[44] PRASANTH, S., HADDAD, Z., SUSILUOTO, J., BRAVERMAN, A., OWHADI, H., HAMZI, B., HRISTOVA-VELEVA, S., and TURK, J. Kernel flows to infer the structure of convective storms from satellite passive microwave observations. 2021 *AGU Fall Meeting Abstracts*, AGU, San Francisco, A55F-1445 (2021)

[45] SUSILUOTO, J., BRAVERMAN, A., BRODRICK, P., HAMZI, B., JOHNSON, M., LAMMINPAA, O., OWHADI, H., SCOVEL, C., TEIXEIRA, J., and TURMON, M. Radiative transfer emulation for hyperspectral imaging retrievals with advanced kernel flows-based Gaussian process emulation. 2021 *AGU Fall Meeting Abstracts*, AGU, San Francisco, NG25A-0506 (2021)

[46] AKIAN, J. L., BONNET, L., OWHADI, H., and SAVIN, É. Learning "best" kernels from data in Gaussian process regression with application to aerodynamics. *Journal of Computational Physics*, **470**, 111595 (2022)

[47] RAISSI, M., YAZDANI, A., and KARNIADAKIS, G. E. Hidden fluid mechanics: learning velocity and pressure fields from flow visualizations. *Science*, **367**, 1026–1030 (2020)