Inferring flow parameters and turbulent configuration with physics-informed data assimilation and spectral nudging

Patricio Clark Di Leoni,[1,*] Andrea Mazzino,[2,†] and Luca Biferale[1,‡]
[1]*Department of Physics and INFN, University of Rome Tor Vergata, Via della Ricerca Scientifica 1, 00133 Rome, Italy*
[2]*Department of Civil, Chemical, and Environmental Engineering and INFN, University of Genova, Genova 16145, Italy*

(Received 9 April 2018; published 11 October 2018)

Inferring physical parameters of turbulent flows by assimilation of data measurements is an open challenge with key applications in meteorology, climate modeling, and astrophysics. Up to now, spectral nudging was applied for empirical data assimilation as a means to improve deterministic and statistical predictability in the presence of a restricted set of field measurements only. Here we explore under which conditions a nudging protocol can be used for two objectives: to unravel the value of the physical flow parameters and to reconstruct large-scale turbulent properties starting from a sparse set of information in space and in time. First, we apply nudging to quantitatively infer the unknown rotation rate and the shear mechanism for turbulent flows. Second, we show that a suitable spectral nudging is able to reconstruct the energy containing scales in rotating turbulence by using a blind setup, i.e., without any input about the external forcing mechanisms acting on the flow. Finally, we discuss the broad potentialities of nudging to other key applications for physics-informed data assimilation in environmental or applied flow configurations.

DOI: 10.1103/PhysRevFluids.3.104604

I. INTRODUCTION

Extracting information from experimental or observational data of fluid flows is a highly challenging task. While in laboratory experiments one can control and/or measure the properties of the system (e.g., viscosity, thermal expansion coefficient, large-scale shear, rotation rate, etc.), this is often impossible when performing observations in the open field, such as for meteorological data taken from the atmosphere or astrophysical data in the sky. Thus, one has to resort to other methods to infer the desired parameters, a task which most of the time is obstructed by the quality of the data at hand. The problem is part of a vaster paradigm that goes under the name of data assimilation and optimal reconstruction, where one is faced with the need to infer the flow parameters or to extrapolate measurements from a sparse subvolume of the flow field to the whole space. The problem is also connected to the need to control and improve predictability for the evolution of chaotic systems by using only a partial set of information about the full trajectory. These problems can be encountered in a wide range of fields, going from atmospherics sciences [1,2], astrophysics [3], optics [4], and medical physics [5]. Several tools have been developed to tackle these challenges. In the context of numerical weather prediction, variational principles and ensemble filters have been developed to fine-tune the parameters in the subgrid models [6–9]. Alternatively,

*patricio.clark@roma2.infn.it
†andrea.mazzino@unige.it
‡luca.biferale@roma2.infn.it

other techniques coupled with Bayesian inference, machine learning, and deep learning have been proposed to estimate the parameter phase space in Reynolds-averaged Navier-Stokes models in engineering problems [10–13]. Also, information theory and statistical mechanics tools such as belief propagation have been used to infer parameters from turbulent flows by looking at the motions of transported particles [14]. Another interesting example is the use of sparse regression methods to discover not only parameters but the actual form of the terms controlling the evolution of a system [15,16].

In this paper we explore a different avenue and show how to infer the physical flow parameters from partial data assimilation by exploiting the equations of motion in a dynamical way, using a technique known as nudging, whose conceptual foundation goes well beyond applications to physics (see [17]). Contrary to the attempts previously mentioned where the modeled flow is usually compared with data by using a cost function, nudging introduces an extra term in the dynamical equation where partial information from field measurements is input and exploited to reconstruct the unmeasured degrees of freedom. Nudging, has been successfully used and developed to input the global circulation model into a regional climate model [18–20]. In this case, due to computational constraints, the global models cannot solve the smallest dynamically active scales so as to have accurate local weather predictions, while the regional models cannot solve for the large planetary cyclonic and anticyclonic circulations. Nudging is applied to match the overlapping scales in each model by forcing the regional model to behave as the global one via a penalty term. Outside numerical weather prediction, nudging has also been rigorously applied to estimate bounds in the data assimilation problem in two-dimensional Navier-Stokes equations [21,22], the three-dimensional Navier-Stokes α model [23], and Rayleigh-Bénard convection [24]. It has also been used to study synchronization in maps and dynamical systems [25]. Here we attempt to benchmark and optimize its performances to the three-dimensional Navier-Stokes equations in the fully developed turbulent regime, characterized by high chaoticity and by a high-dimensional strange attractor.

We implement a spectral nudging technique with two aims. First, we show how to use nudging as a physics-informed tool to accurately infer key flow parameters, e.g. the rotation rate or the large-scale stirring mechanism, from a limited subset of data sparsely measured in time and in Fourier space. Second, we show that the same technique can be used to learn the global physical turbulent configuration. We do this by using the nudged equations to reconstruct in space the large-scale energy distribution of rotating turbulence under the presence of a split energy cascade and without inputting into the algorithm any information about the external forcing mechanism and about the intensity of the rotation rate. Nudging is thus presented as a general data-driven algorithm to learn from sparse measurements in a dynamical way and with a broad range of applications. Finally, we discuss a series of open challenges to adapt and extend the application of nudging to other turbulent flow configurations using either Eulerian or Lagrangian field measurements and in different domains.

II. NUDGING TECHNIQUE

As said, nudging means to gently convince a numerical flow to evolve as close as possible to a reference set supposing it has only partial measurements or observations of the latter [18–20]. The idea is to use the equation of motion to perform an optimal data and flow-parameter assimilation in the interval of time $t \in (0, t')$ and in the whole fluid volume. Suppose we have a reference three-dimensional turbulent flow $\mathbf{u}_{\rm ref}(\mathbf{x}, t)$ evolving under the action of a set of external forces $\mathcal{F}[\mathbf{u}_{\rm ref}, \mathcal{V}_{\rm ref}]$ parametrized by a set of physical coefficients $\mathcal{V}_{\rm ref} = (\Omega_{\rm ref}, \mathcal{S}_{\rm ref}, \ell_{\rm ref}, \Delta T_{\rm ref}, \ldots)$, where we denote by $\Omega_{\rm ref}$ the rotation rate, by $\mathcal{S}_{\rm ref}$ the amplitude of a large-scale shear with the typical length scale $\ell_{\rm ref}$, and by $\Delta T_{\rm ref}$ the temperature difference across the volume. Suppose that we have access to the measurements of the reference velocity field $\mathbf{u}_{\rm ref}$ on a limited set of M anemometers placed in $\mathbf{x}_j$ with $j = 1, \ldots, M$ that record the flow properties at N time instants t_n with $n = 1, \ldots, N$, i.e., we control $\mathbf{u}_{\rm ref}$ in a given subdomain of the whole space-time $(3 + 1)$ volume

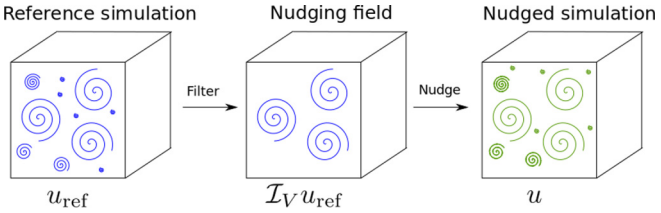

FIG. 1. Diagram showing the setup of our numerical experiments. First, a reference simulation is performed (left). Second, a subset of data is filtered out of the reference field by keeping only data on a given subset of points in space and instant in times (middle). Third, we interpolate in time the input partial information and use it to nudge the evolution of a new field to reconstruct the missing data and to infer the correct physics parameters (right).

only. The idea behind nudging is to evolve an independent three-dimensional incompressible Navier-Stokes equations with an initially educated guess for the set of parameters $\mathcal{V}$ and imposing a penalty whenever the flow field does not reproduce the input velocity values of the reference field in the space-time domain $V = (\mathbf{x}_j, t_n)$,

$$\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u} \cdot \nabla \boldsymbol{u} = -\nabla p + \nu \nabla^2 \boldsymbol{u} + \mathcal{F}[\mathbf{u}, \mathcal{V}] - \alpha \mathcal{I}_V (\boldsymbol{u} - \boldsymbol{u}_{\mathrm{ref}}), \tag{1}$$

where ν is the viscosity, p is the pressure that ensures the incompressibility condition, $\mathcal{I}_V$ is a dimensionless linear projector operator given by the characteristic function of the set V, and α is a parameter that controls the intensity imposed by the nudging control and has units of frequency. In its crudest form, $\mathcal{I}_V$ is equal to 1 for $(\mathbf{x}, t) \in V$ and 0 otherwise. The simplest and most common improvement is to linearly interpolate the different measured snapshots between each time t_n and t_{n+1}. So when included in (1), $\mathbf{u}_{\mathrm{ref}}$ will always be assumed to be piecewise differentiable in time with a characteristic interpolation window τ. In this way the operator $\mathcal{I}_V$ is only acting on the spatial part of the fields. The whole protocol is sketched in Fig. 1. It is important to realize that, in our application, we do not even require one to know the exact way the system is forced, i.e., we do not impose $\mathcal{V} = \mathcal{V}_{\mathrm{ref}}$ and the only *a priori* information that we provide is inside the partial measurements of the reference field. Clearly, the success of the reconstruction will depend on the amount of information provided (how many measurements in space and in time), on its quality (where and what we measure), and on the intensity of the penalization term α. Notice that, because of potential stiffness and truncation effects arising when α is big, it is not *a priori* obvious that taking large α is the best choice. It is intuitive to imagine that in some cases it might be better to allow for a larger error in some measuring stations to allow the field to be closer to the target globally.

III. SETUP OF THE NUMERICAL SPECTRAL NUDGING EXPERIMENT

We start first by restricting the discussion to the case when the set of external parameters is given by the intensity of the Coriolis force due to the presence of a rotation Ω in the vertical direction and of an external stirring mechanism $\boldsymbol{\mathcal{S}}$,

$$\mathcal{F}[\mathbf{u}, \mathcal{V}] = 2\Omega \hat{\mathbf{z}} \times \mathbf{u}(\mathbf{x}, t) + \boldsymbol{\mathcal{S}}(\mathbf{x}), \tag{2}$$

where $\boldsymbol{\mathcal{S}}$ is a randomly generated, quenched in time, isotropic field with support on wave numbers with amplitudes $k \in [k_{f1}, k_{f2}]$ whose Fourier coefficients are given by $\hat{\boldsymbol{\mathcal{S}}}(\mathbf{k}) = \mathcal{S} k^{-7/2} e^{i\theta_{\mathbf{k}}}$, where $\theta_{\mathbf{k}}$ are the random phases. In the rest of this paper we will address the most ideal case when the information is supplied in Fourier space, i.e., we imagine having a periodic array of measurement stations that allow us to reconstruct the reference flow configuration in a given range of nudged wave numbers $k_0 < k < k_1$. In this case, the $\mathcal{I}_V$ operator reduces to a bandpass Fourier filter of the

TABLE I. Parameters used in the different numerical experiments. INFER1 is the setup for the Ω scan, INFER2 for the $\mathcal{S}$ scan, and PHYS1 for the inverse cascade experiment. The values listed are the total kinetic energy $E_{\rm kin} = \frac{1}{2}\langle|\mathbf{u}|^2\rangle$, the eddy turnover time $T = L/(2E_{\rm kin})^{1/2}$ with $L = 2\pi$ the largest scale in the flow (same for all simulations), the viscosity ν, the Reynolds number ${\rm Re} = L(2E_{\rm kin})^{1/2}/\nu$, the forcing intensity of the reference simulation $\mathcal{S}_{\rm ref}$, the band of forced wave numbers in the reference simulation $[k_{f1}, k_{f2}]$, the rotation frequency of the reference simulation $\Omega_{\rm ref}$, and the band of nudged wave numbers $[k_0, k_1]$. The number of grid points $N^3_{\rm grid} = 256^3$, the time step of the simulations $dt = 0.001$, the nudging intensity $\alpha = 10$, and the temporal interpolation window of the nudging field $\tau = 0.1$ are the same for all simulations. The box length L, the temporal time step dt, the resolution $N^3_{\rm grid}$, and the viscosity ν are the same for the reference and the nudged simulations in each set. The kinetic energy and Reynolds numbers are given for the reference run of each set; the nudged ones have very similar values.

Setup	$E_{\rm kin}$	T	ν	Re	$\mathcal{S}_{\rm ref}$	$[k_{f1}, k_{f2}]$	$\Omega_{\rm ref}$	$[k_0, k_1]$
INFER1	1.84	3.28	0.002	6030	0.005	[1,2]	2	[1,4]
INFER2	1.20	4.06	0.0025	4900	0.02	[1,2]	0	[1,4]
PHYS1	0.0012	128	0.002	150	0.004	[10,11]	20	[8,20]

form

$$\mathcal{I}_V \mathbf{u} = \sum_{k_0<|\mathbf{k}|<k_1} \hat{\mathbf{u}}(\mathbf{k}, t) \exp(i\mathbf{k}\cdot\mathbf{x}) \tag{3}$$

that projects the velocity field on the window of nudged Fourier modes.

We implement the whole protocol as follows. First, we numerically produce a full space-time evolution of the whole $\mathbf{u}_{\rm ref}$ field in an interval $t \in (0, T_{\rm tot})$ by solving the Navier-Stokes equations with a reference rotation rate $\Omega_{\rm ref}$ and a given intensity of the shear $\mathcal{S}_{\rm ref}$ [i.e., Eq. (1) with $\alpha = 0$]. The values of $\Omega_{\rm ref}$ and $\mathcal{S}_{\rm ref}$ (and also ν, which is the same for both the reference and the nudged simulations) are given in Table I. All reference simulations start from rest and are allowed to reach stationary states ($t = 0$ denotes the start of the stationary states). Second, we extract the inputting field in a subset of discrete times $t_n = n\tau$ with τ chosen as a fraction of the characteristic eddy turnover time of the flow (see Table I). Third, we define the nudging field (3) by a linear interpolation between t_n and t_{n+1} for all intervals. The initial condition used for all nudged simulation is just the first extracted input field (i.e., the field at $t = t_0$) with all the modes outside the nudging region filtered out. All simulations have been performed with a parallel pseudospectral code. The code uses a two-step Adams-Bashforth scheme for the time integration, the 2/3 rule for dealiasing, and periodic boundary conditions in all three directions. In the following we will analyze three different nudging protocols. The first two cases are about simulations made to infer the physical flow parameters $\Omega_{\rm ref}$ and $\mathcal{S}_{\rm ref}$ (called INFER1 and INFER2 in the following; see also Table I for details). The third case is about the reconstruction of the large-scale coherent structures and it is called PHYS1. Numerical details for all setups can be found in Table I. The value of τ is such that it is smaller than the decorrelation time of the fastest nudged mode, while α was taken as $1/\tau$; these choices follow common practices [26]. A comprehensive report about the performance of nudging at changing α and τ for fully developed homogeneous and isotropic turbulent flow is beyond the scope of this paper.

IV. INFERRING PHYSICAL PARAMETERS IN ROTATING TURBULENCE

We start by asking how to guess the exact value of the rotation rate $\Omega_{\rm ref}$ without any *a priori* knowledge of its value. To give an initial idea about the applications of nudging, in Figs. 2(a)–2(d) we show a series of two-dimensional (2D) slices of the vorticity field in the direction parallel to the rotation axis for the reference simulation [Fig. 2(a)] and for three different nudged simulations [Figs. 2(b)–2(d)], two with wrong rotation rates $\Omega = 0$ and $\Omega = 2\Omega_{\rm ref}$ and one with the correct

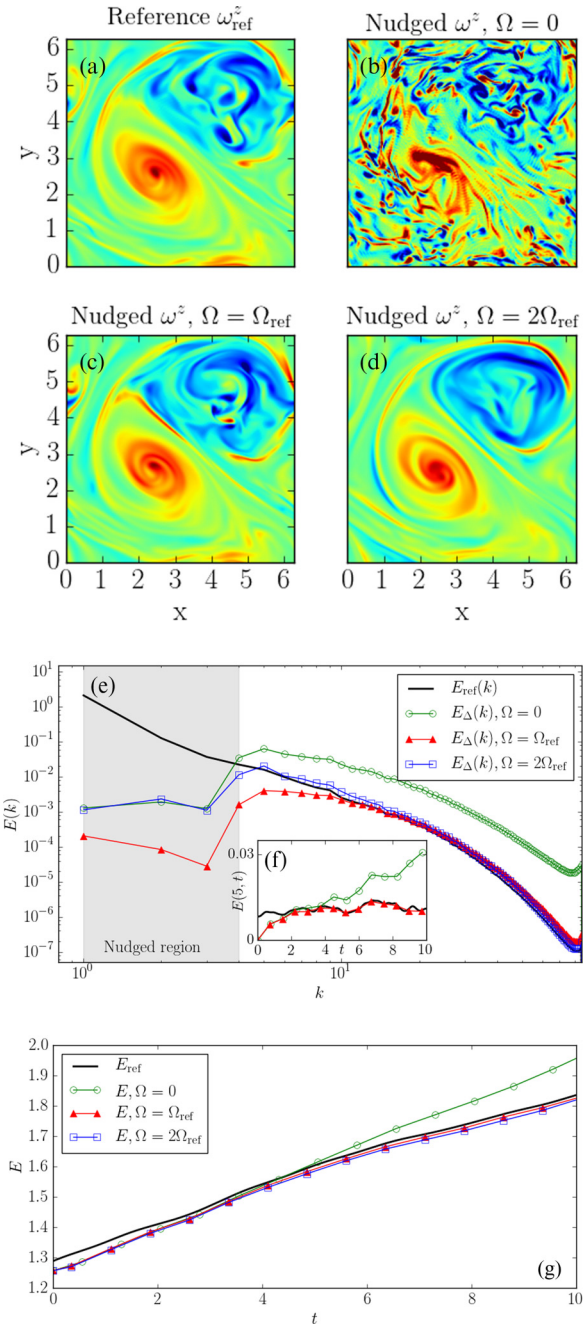

FIG. 2. Nudging with different rotation rates. Simulations from setup INFER1 (see Table I). Two-dimensional slices of the vorticity field $\boldsymbol{\omega} = \boldsymbol{\nabla} \times \mathbf{u}$ in the direction parallel to the rotation axis are shown for (a) the reference simulation with a rotation frequency $\Omega_{\rm ref}$ and three nudged simulations performed with (b) $\Omega = 0$, (c) $\Omega = \Omega_{\rm ref}$, and (d) $\Omega = 2\Omega_{\rm ref}$. (e) Energy spectra of the reference simulation compared with error spectra $E_\Delta(k)$ [see Eq. (4)] for different values of the rotation frequency Ω. All spectra were computed at the same instants of time. The shaded gray area indicates the modes where the nudging is acting. (f) Time evolution of $E(k, t)$ for $k = 5$ for $\Omega = \Omega_{\rm ref}$ and $\Omega = 0$, compared to the reference data. (g) Evolution of total energy for the reference field and for the three nudged simulations at changing Ω.

value $\Omega = \Omega_{\text{ref}}$. Furthermore, in this set of simulations we took $\mathcal{S} = 0$, i.e., we assume to not know the forcing mechanism (all simulations are from setup INFER1 shown in Table I). All snapshots were taken at the same instant in time. Comparing the four panels, it is clear that the simulation nudged with the correct rotation rate [Fig. 2(c)] does reconstruct the reference flow [Fig. 2(a)] much better than the other two [Figs. 2(b) and 2(d)]. It is also worth pointing out that the standard deviation of the vorticity fields is recovered when rotation is present, with the values being around 2.8 for the reference and the simulations of both Figs. 2(c) and 2(d), but this is not the case in the absence of rotation [Fig. 2(c)], where the standard deviation takes a value around 5.4. All fields have zero mean by construction. These qualitative results already provide a first glance at the two main points we make: (i) Spectral nudging works well also for fully turbulent 3D flows, as it reproduces nontrivial features with high accuracy, and (ii) by optimizing the reconstruction properties, one can infer the unknown flow parameters of the nudging flow. It is worth noticing that the percentage of nudged modes is very small, of the order of $\mathcal{N}_{\text{nudged}} \sim 1 \times 10^{-4}$, as we are nudging up to $k = 4$, while the maximum possible wave number in this simulation is $k = 85$. The nudged modes are the ones containing the largest amount of energy, but the flow is not completely determined by their evolution, as many more scales should be controlled in order to achieve this [27]. This fact is clear when looking at the error spectra in Fig. 2, where the error in the unnudged scales is of the order of the energy at those scales even though the large-scale reconstruction is very good, meaning the unnudged scales are not slaved to the energy containing modes. Some synchronization of the small scales is nonetheless present, especially for the case with $\Omega = \Omega_{\text{ref}}$. Understanding how much one needs to nudge in order to fully control a turbulent flow is an open question that is left for future work.

In order to control the performance of the nudging protocol in quantitative terms and scale by scale, we introduce a field given by the difference among the exact input and the one reconstructed via (1), $\Delta\mathbf{u} = \mathbf{u} - \mathbf{u}_{\text{ref}}$, and we study its spectral properties:

$$E_{\Delta}(k,t) = \frac{1}{2} \sum_{k \leqslant |\mathbf{k}| < k+1} |\hat{\boldsymbol{u}}(\mathbf{k},t) - \hat{\boldsymbol{u}}_{\text{ref}}(\mathbf{k},t)|^2. \tag{4}$$

Clearly, the smaller the spectrum $E_{\Delta}(k)$, the better the reconstruction. This spectrum will be referred to as the error spectrum.

In Fig. 2(e) we show three different curves for $E_{\Delta}(k,t)$ obtained by averaging over all times when we provide the information t_n and for the three different values of the rotation rate $\Omega = 0, \Omega_{\text{ref}}, 2\Omega_{\text{ref}}$, already discussed in Figs. 2(a)–2(d), together with the spectrum of the reference field $E_{\text{ref}}(k) = \sum_{k \leqslant |\mathbf{k}| < k+1} |\hat{\boldsymbol{u}}_{\text{ref}}(\mathbf{k},t)|^2$, averaged on the same set of times. In the figure, the set of nudged wave numbers is denoted by the gray area. From Fig. 2(e) it is clear that the optimal nudging is obtained when $\Omega = \Omega_{\text{ref}}$ is used in (1), as revealed from the scale-by-scale nudging error $E_{\Delta}(k)$ that becomes much smaller than $E_{\text{ref}}(k)$ for $k \in (k_0, k_1)$. In all cases, there is a dip in the error spectra at $k = 3$, as this is the first scale at which the forcing is not present in the reference flow, so the nudging is able to do a better job reconstructing the data. At $k = 4$ the error spectra increase again, mainly because some unnudged modes are integrated when calculating the spectra at this wave number. For $\Omega = \Omega_{\text{ref}}$, the scale-by-scale error stays smaller than the reference spectrum up to $k \sim 10$, suggesting the ability to assimilate data outside the set of nudged degrees of freedom also. This latter fact is also confirmed by the inset [Fig. 2(f)], where we show the temporal evolution of $E_{\text{ref}}(k,t)$ for an unnudged wave number $k = 5$ compared with the spectra of the reconstructed field evolving with $\Omega = 0$ and $\Omega = \Omega_{\text{ref}}$. In this experiment we started the nudged simulations from zero velocity. As one can see, after a short transient, only the field evolving with the correct Ω rate is indeed able to synchronize with the time evolution of the inputting data. Finally, we also show the evolution of the total energy in Fig. 2(g) for the same simulations. While the case with $\Omega = 0$ is easy to discern, the other two are too close to tell which one produces a better reconstruction of the flow. This indicates that comparing averaged quantities (such as the total energy) may not be the most precise way to determine the value of a parameter.

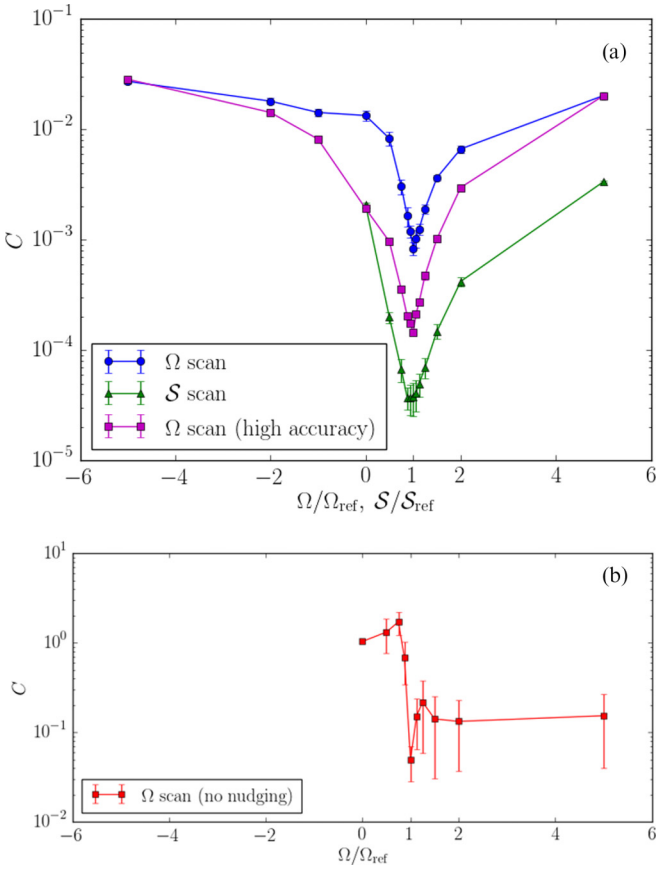

FIG. 3. (a) Value of the mean error committed to reconstruct the reference field in the nudged window C for two different scans of the phase-space parameters. Blue circles show the case with a fixed stirring mechanism and changing rotation rate Ω (setup INFER1 in Table I). Green triangles show the case with fixed rotation rate and changing intensity of the stirring parameter $\mathcal{S}$ (setup INFER2 in Table I). Magenta squares show a further scan for Ω following INFER1 but nudging all wave numbers up to $k = 10$. In all cases a clear depth is measured only when the scanning values correspond to the ones used for the reference data $\Omega_{\rm ref}$ and $\mathcal{S}_{\rm ref}$, respectively. Error bars for each data point were calculated by measuring the standard deviation of C. (b) Scan of Ω performed without nudging but adding the forcing term (i.e., same as INFER1 but with $\alpha = 0$ and $\mathcal{S} = \mathcal{S}_{\rm ref}$).

To be more quantitative about the sensitivity to infer the unknown rotation rate, we have performed also a detailed scan of Ω values around $\Omega_{\rm ref}$. In Fig. 3(a) we show the performance of the nudging reconstruction by plotting the value of the spectrum $E_\Delta(k)$ as a function of Ω and averaged in time and in the nudged window

$$C = \frac{1}{NK} \sum_{n=1}^{N} \int_{k_0}^{k_1} dk\, E_\Delta(k, t_n), \tag{5}$$

where t_n are the instants in time where we have measurements and $K = \int_{k_0}^{k_1} dk\, E_{\rm ref}(k)$ is a normalization factor. Notice that C is defined using information about the nudging data only, i.e., the filtered reference field at the specific times when the information is provided. In contrast, $E_\Delta(k)$ needs the whole $\mathbf{u}_{\rm ref}$, which in most practical applications would not be available, but which we can nevertheless access in our numerical experiment.

From Fig. 3(a) the existence of a minimum in the error when evolving (1) with $\Omega \sim \Omega_{\rm ref}$ is clear. Furthermore, we can determine the correct value of Ω with a 6.25% error. The error is calculated by looking at which values the error bars for C overlap. We performed another experiment (setup INFER2 in Table I) to test if the intensity $\mathcal{S}$ of mechanical forcing of the reference simulation could also be discovered with our nudging protocol. In this experiment a reference simulation with $\Omega_{\rm ref} = 0$ and $\mathcal{S}_{\rm ref} = 0.02$ was produced and used to extract the nudging fields (see Table I for details). In Fig. 3(a) we show that the protocol is able to infer the intensity of the stirring mechanism also, with a clear minimum of the error (5) in the proximity of $\mathcal{S} \sim \mathcal{S}_{\rm ref}$. In this case, the correct value of $\mathcal{S}$ can be pinpointed with a 12.5% error. A third experiment, following INFER1 but nudging more wave numbers (so using more information from the reference as well) is shown. Here all wave numbers up to $k = 10$ where nudged. By doing this we can reduce the error in the estimation of Ω to 3.125%. All numerical experiments show that spectral nudging can be used in a physics-informed way to fit parameters to data and thus extract information from it. Furthermore, in setup INFER1, where no information about the external stirring mechanism is used, performing a one-dimensional scan (i.e., varying only the rotation rate) works well. Having said this, we cannot conclude that this must be the case for a generic search in a multidimensional phase space, where the only systematic way to proceed would be to adopt a local gradient-descent algorithm.

A similar scan was performed for the rotation rate but without using the nudging (i.e., $\alpha = 0$). In this case, the forcing term was also added (with $\mathcal{S} = \mathcal{S}_{\rm ref}$); otherwise there would be no energy injection mechanism present. All other parameters are the same as for setup INFER1. The results are shown in Fig. 3(b). It is clear that obtaining an accurate value of Ω from scan is very difficult because even though a minimum is readily seen, the error bars of several data points close to it overlap. So while running simulations with different parameter values and performing *a posteriori* analysis in order to infer the desired information is possible, our results suggest that using nudging greatly improves the sensitivity and accuracy of the search.

V. INFERRING THE LARGE-SCALE VELOCITY DISTRIBUTION WITHOUT INPUT ROTATION

In this section we describe how to use nudging to infer, under some circumstances, the entire set of large-scale physical flow structures of the reference data without detailed knowledge of the forces acting on flow. To test this idea we performed an experiment by using a turbulent flow under rotation and in the presence of an inverse energy cascade. It is well known that if rotation is strong enough and energy is injected at large wave numbers the flow undergoes a transition from a direct to a split turbulent energy cascade, accumulating kinetic energy and producing a nontrivial cyclonic distribution of vortices at larger and larger scales [28–30]. This regime does not occur naturally in homogeneous isotropic three-dimensional turbulence [31], but it is argued to be important in many geophysical setups in oceans [32,33] and the atmosphere [34]. Here we show how a suitable nudging strategy is indeed able to reconstruct the inverse energy cascade even in the absence of any explicit rotation term in the nudged equations (1), provided the $\mathbf{u}_{\rm ref}$ is inputting information around the injection scale. To do this we use a rotating turbulent flow forced at $k_f = 10$ and with $\Omega_{\rm ref} = 20$ and $\mathcal{S}_{\rm ref} = 0.004$ as a reference (setup PHYS1 in Table I) where an inverse energy cascade develops. We then evolve (1) without any rotation and any external forcing,

$$\mathcal{F}[\mathbf{u}_{\rm ref}, \mathcal{V}_{\rm ref}] = 0;$$

in this way we are completely ignorant about the physics we want to reproduce. In Fig. 4 we show that by nudging in the region around the injection mechanism, the energy spectrum of the reference simulation is well reproduced by the nudged simulation even, and in particular, in the inverse energy cascade range. The presence of a strong peak around the forced wave number is typical of systems where an inverse cascade is present, as this is a slow and inefficient transfer mechanism [28–30,35]. Even though the only information we input is the nudging filtered field, the nudged evolution is able to reconstruct the inverse cascade and the correct spectrum slope even for scales much smaller than the ones where we nudge. It is remarkable how the spectrum error $E_\Delta(k)$ is small also for modes

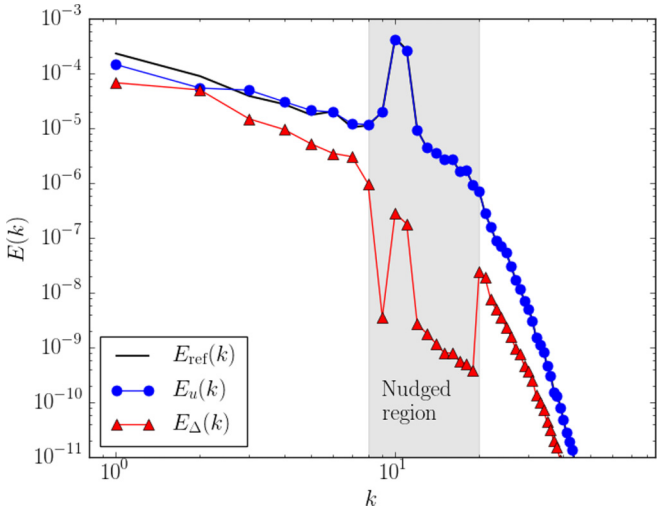

FIG. 4. Nudging for the case of rotating turbulence in the inverse energy cascade regime. Simulations are from setup PHYS1 (see Table I). The nudged window is given by the gray area between $k_0 = 8$ and $k_1 = 20$. The reference spectrum $E_{\rm ref}$ and the nudged spectrum E_u almost coincide for $k > 8$, making it hard to discern between the two. Both the intensity of the forcing $\mathcal{S}$ and the rotation rate Ω are zero in the nudged simulation, so all energy injection and anisotropic effects are coming from the nudging term. Notice the strongly reduced error spectrum $E_\Delta(k)$ for a large set of wave numbers, indicating an optimal reconstruction quality.

outside the nudging window $k < k_0$ and $k > k_1$, indicating the presence of strong nonlocal spectral correlation in the split-energy cascade mechanism which is fully reconstructed by our protocol.

To go beyond spectral properties and to check the ability to reconstruct the large-scale coherent structures in the rotating flow, we plot in Fig. 5 the probability density functions (PDFs) of the space-dependent kinetic energy for the reference simulation $|\mathbf{u}_{\rm ref}|^2/2$, the nudged simulation $|\mathbf{u}|^2/2$,

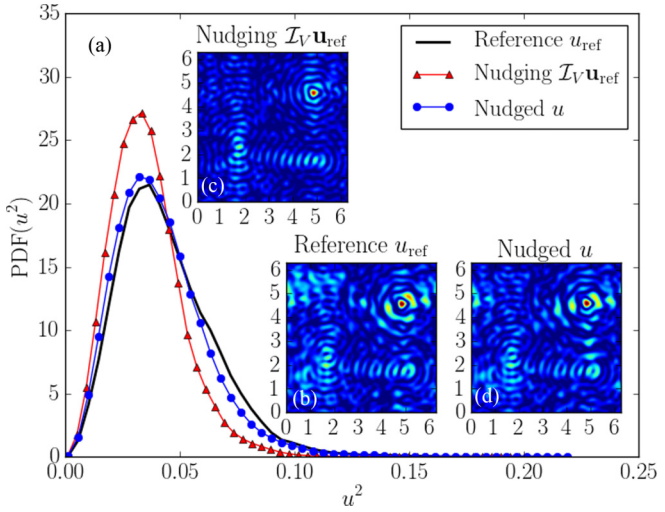

FIG. 5. (a) Probability density functions of the pointwise kinetic energy for the reference simulation $|\mathbf{u}_{\rm ref}(\mathbf{x})|^2$ (solid black line), the nudged simulation $|\mathbf{u}(\mathbf{x})|^2$ (circles), and the nudging input field $|\mathcal{I}_V\mathbf{u}_{\rm ref}(\mathbf{x})|^2$ (triangles) for the inverse cascade experiment (setup PHYS1 in Table I). (b)–(d) Two-dimensional slices of planes perpendicular to the rotation axis of the absolute velocity fields.

and the nudging field $|\mathcal{I}_V \mathbf{u}_{\rm ref}|^2/2$. As one can see, the reconstructed field has a PDF very close to the reference case, even if the nudging input field does not. In the same figure we also show 2D slices of the absolute velocity fields in planes perpendicular to the rotation axis for the three fields as before. As one can see, the nudged simulation [Fig. 5(d)] is able to extrapolate the unknown large-scale reference flow structures extremely well [Fig. 5(b)], for a case where the nudged inputting data do not contain any information about those scales [Fig. 5(c)]. The apparent patterns seen in these visualizations are a product of the strong forcing present in the system acting around $k = 10$.

VI. CONCLUSION

Spectral nudging is a physics-informed technique commonly used to guide the evolution of chaotic dynamical systems inputting measured data. Giving examples for both isotropic and rotating 3D turbulence, we have shown how this technique can be efficiently used to infer both the physical parameters of the external stirring forces and the large-scale velocity distribution for the inverse energy cascade regime, typical of strongly rotating turbulent flows. The method can be further improved and optimized by using different nudging parameters for different degrees of freedoms, e.g., by changing α and τ with $\mathbf{k}$. A detailed study of nudging performances for homogeneous and isotropic turbulence at different Reynolds numbers, at different nudging windows, and at changing spatial locations of the measurements stations is left for future work.

Other strategies used to estimate parameters, such as variational methods [36] or ensemble based methods [7–9], require the need to postulate an error correlation matrix and make assumptions about the behavior of the errors and deviations, need to use linearized models (for variational methods), or are based on minimizing complicated functions (again for variational methods). Nudging-based strategies require one to perform several forward simulations, similar to the ensemble-based method. One advantage other methods have compared to nudging is the ease to incorporate information on observables (such as precipitation, for example) and not just state variables (such as the velocity field, as was used here). Interestingly, variational data-assimilation schemes have been exploited to determine vectors of optimal nudging coefficients [37]. Here we reversed the point of view: Given the coefficients α and τ, we employed nudging to estimate the physical flow parameters. Finally, the method is also general and extendable to other problems, opening the route to applications for parameter extrapolation to a vast set of hydrodynamic situations including the most promising cases such as (i) optimizing subgrid-scale models in large-eddy simulations by inferring parameters against data extracted from either observation or benchmark direct numerical simulations, (ii) large-scale turbulent transport to determine eddy viscosity and eddy diffusivity [38,39], (iii) the identification of ambient air sources and the quantification of their contribution to pollution levels (the so-called source apportionment problem) [40], (iv) partial field reconstruction using advanced lidar systems [41] to reveal the free parameters characterizing the atmospheric boundary layer, and (v) correction of velocity fields in ocean circulation models with Lagrangian data (e.g., from drifting buoys) [42,43] and/or other sources including high-frequency radar data [44].

ACKNOWLEDGMENT

The authors acknowledge funding from the European Research Council under the European Community's Seventh Framework Program, ERC Grant Agreement No. 339032.

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, Wireless sensor networks: A survey, Comput. Netw. **38**, 393 (2002).

[2] J. K. Hart and K. Martinez, Environmental sensor networks: A revolution in the earth system science? Earth-Sci. Rev. **78**, 177 (2017).

[3] H. S. Fu, A. Vaivads, Y. V. Khotyaintsev, V. Olshevsky, M. André, J. B. Cao, S. Y. Huang, A. Retinó, and G. Lapenta, How to find magnetic nulls and reconstruct field topology with MMS data? J. Geophys. Res.: Space Phys. **120**, 2015JA021082 (2015).
[4] P. Carpeggiani, M. Reduzzi, A. Comby, H. Ahmadi, S. Kühn, F. Calegari, M. Nisoli, F. Frassetto, L. Poletto, D. Hoff, J. Ullrich, C. D. Schröter, R. Moshammer, G. G. Paulus, and G. Sansone, Vectorial optical field reconstruction by attosecond spatial interferometry, Nat. Photon. **11**, 383 (2017).
[5] J. Busch, D. Giese, L. Wissmann, and S. Kozerke, Reconstruction of divergence-free velocity fields from cine 3D phase-contrast flow measurements, Magn. Reson. Med. **69**, 200 (2013).
[6] E. Kalnay, *Atmospheric Modeling, Data Assimilation and Predictability* (Cambridge University Press, Cambridge, 2003).
[7] J. L. Anderson and S. L. Anderson, A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts, Mon. Weather Rev. **127**, 2741 (1999).
[8] J. L. Anderson, An ensemble adjustment Kalman filter for data assimilation, Mon. Weather Rev. **129**, 2884 (2001).
[9] J. J. Ruiz, M. Pulido, and T. Miyoshi, Estimating model parameters with ensemble-based data assimilation: A review, J. Meteorol. Soc. Jpn. Ser. II **91**, 79 (2013).
[10] M. C. Kennedy and A. O'Hagan, Bayesian calibration of computer models, J. R. Stat. Soc. Ser. B **63**, 425 (2001).
[11] H. Xiao, J. L. Wu, J. X. Wang, R. Sun, and C. J. Roy, Quantifying and reducing model-form uncertainties in Reynolds-averaged Navier-Stokes simulations: A data-driven, physics-informed Bayesian approach, J. Comput. Phys. **324**, 115 (2016).
[12] E. J. Parish and K. Duraisamy, A paradigm for data-driven predictive modeling using field inversion and machine learning, J. Comput. Phys. **305**, 758 (2016).
[13] J. Ling, A. Kurzawski, and J. Templeton, Reynolds averaged turbulence modeling using deep neural networks with embedded invariance, J. Fluid Mech. **807**, 155 (2016).
[14] M. Chertkov, L. Kroc, F. Krzakala, M. Vergassola, and L. Zdeborová, Inference in particle tracking experiments by passing messages between images, Proc. Natl. Acad. Sci. USA **107**, 7663 (2010).
[15] S. L. Brunton, J. L. Proctor, and J. N. Kutz, Discovering governing equations from data by sparse identification of nonlinear dynamical systems, Proc. Natl. Acad. Sci. USA **113**, 3932 (2016).
[16] S. H. Rudy, S. L. Brunton, J. L. Proctor, and J. N. Kutz, Data-driven discovery of partial differential equations, Sci. Adv. **3**, e1602614 (2017).
[17] R. E. Thaler, Nobel lecture, From cashews to nudges: The evolution of behavioral economics, The Sveriges Riksbank Prize in Economic Sciences (2017), available at https://www.nobelprize.org/prizes/economics/2017/thaler/lecture/
[18] K. M. Waldron, J. Paegle, and J. D. Horel, Sensitivity of a spectrally filtered and nudged limited-area model to outer model options, Mon. Weather Rev. **124**, 529 (1996).
[19] H. von Storch, H. Langenberg, and F. Feser, A spectral nudging technique for dynamical downscaling purposes, Mon. Weather Rev. **128**, 3664 (2000).
[20] G. Miguez-Macho, G. L. Stenchikov, and A. Robock, Spectral nudging to eliminate the effects of domain position and geometry in regional climate model simulations, J. Geophys. Res. Atmos. **109**, D13104 (2004).
[21] A. Farhat, E. Lunasin, and E. S. Titi, Abridged continuous data assimilation for the 2D Navier-Stokes equations utilizing measurements of only one component of the velocity field, J. Math. Fluid Mech. **18**, 1 (2016).
[22] M. Gesho, E. Olson, and E. S. Titi, A computational study of a data assimilation algorithm for the two-dimensional Navier-Stokes equations, Commun. Comput. Phys. **19**, 1094 (2016).
[23] D. A. F. Albanez, H. J. Nussenzveig Lopes, and E. S. Titi, Continuous data assimilation for the three-dimensional Navier-Stokes-α model, Asymptotic Anal. **97**, 139 (2016).
[24] A. Farhat, H. Johnston, M. Jolly, and E. S. Titi, Assimilation of nearly turbulent Rayleigh-Bénard flow through vorticity or local circulation measurements: A computational study, J. Sci. Comput., doi:10.1007/s10915-018-0686-x.

[25] D. Pazó, J. M. López, R. Gallego, and M. A. Rodríguez, Synchronizing spatio-temporal chaos with imperfect models: A stochastic surface growth picture, Chaos **24**, 043115 (2014).
[26] H. Omrani, P. Drobinski, and T. Dubos, Spectral nudging in regional climate modeling: How strongly should we nudge? Q. J. R. Meteorol. Soc. **138**, 1808 (2012).
[27] C. C. Lalescu, C. Meneveau, and G. L. Eyink, Synchronization of Chaos in Fully Developed Turbulence, Phys. Rev. Lett. **110**, 084102 (2013).
[28] L. M. Smith, J. R. Chasnov, and F. Waleffe, Crossover from Two- to Three-Dimensional Turbulence, Phys. Rev. Lett. **77**, 2467 (1996).
[29] P. D. Mininni and A. Pouquet, Helicity cascades in rotating turbulence, Phys. Rev. E **79**, 026304 (2009).
[30] P. D. Mininni, A. Alexakis, and A. Pouquet, Scale interactions and scaling laws in rotating flows at moderate Rossby numbers and large Reynolds numbers, Phys. Fluids **21**, 015108 (2009).
[31] L. Biferale, S. Musacchio, and F. Toschi, Split energy-helicity cascades in three-dimensional homogeneous and isotropic turbulence, J. Fluid Mech. **730**, 309 (2013).
[32] R. B. Scott and F. Wang, Direct evidence of an oceanic inverse kinetic energy cascade from satellite altimetry, J. Phys. Oceanogr. **35**, 1650 (2005).
[33] R. Corrado, G. Lacorata, L. Palatella, R. Santoleri, and E. Zambianchi, General characteristics of relative dispersion in the ocean, Sci. Rep. **7**, 46291 (2017).
[34] G. Lacorata, E. Aurell, B. Legras, and A. Vulpiani, Evidence for a $k^{-5/3}$ spectrum from the EOLE Lagrangian balloons in the low stratosphere, J. Atmos. Sci. **61**, 2936 (2004).
[35] L. Biferale, F. Bonaccorso, I. M. Mazzitelli, M. A. T. van Hinsberg, A. S. Lanotte, S. Musacchio, P. Perlekar, and F. Toschi, Coherent Structures and Extreme Events in Rotating Multiphase Turbulent Flows, Phys. Rev. X **6**, 041036 (2016).
[36] I. M. Navon, Practical and theoretical aspects of adjoint parameter estimation and identifiability in meteorology and oceanography, Dyn. Atmos. Oceans **27**, 55 (1998).
[37] X. Zou, I. M. Navon, and F. X. Ledimet, An optimal nudging data assimilation scheme using parameter estimation, Q. J. R. Meteorol. Soc. **118**, 1163 (1992).
[38] L. Yu and J. J. O'Brien, Variational estimation of the wind stress drag coefficient and the oceanic eddy viscosity profile, J. Phys. Oceanogr. **21**, 709 (1991).
[39] C.-L. Lin, T. Chai, and J. Sun, Retrieval of flow structures in a convective boundary layer using an adjoint model: Identical twin experiments, J. Atmos. Sci. **58**, 1767 (2001).
[40] M. C. Bove, P. Brotto, F. Cassola, E. Cuccia, D. Massabò, A. Mazzino, A. Piazzalunga, and P. Prati, An integrated PM2.5 source apportionment study: Positive matrix factorisation vs. the chemical transport model CAMx, Atmos. Environ. **94**, 274 (2014).
[41] D. I. Cooper, W. E. Eichinger, R. E. Ecke, J. C. Y. Kao, J. M. Reisner, and L. L. Tellier, Initial investigations of microscale cellular convection in an equatorial marine atmospheric boundary layer revealed by lidar, Geophys. Res. Lett. **24**, 45 (1997).
[42] V. Taillandier, A. Griffa, and A. Molcard, A variational approach for the reconstruction of regional scale Eulerian velocity fields from Lagrangian data, Ocean Model. **13**, 1 (2006).
[43] V. Taillandier, A. Griffa, P.-M. Poulain, and K. Béranger, Assimilation of Argo float positions in the north western Mediterranean Sea and impact on ocean circulation simulations, Geophys. Res. Lett. **33**, L11604 (2006).
[44] M. Berta, L. Bellomo, M. G. Magaldi, A. Griffa, A. Molcard, J. Marmain, M. Borghini, and V. Taillandier, Estimating Lagrangian transport blending drifters with HF radar data and models: Results from the TOSCA experiment in the Ligurian current (north western Mediterranean Sea), Prog. Oceanogr. **128**, 15 (2014).