Optimal eddy viscosity in closure models for two-dimensional turbulent flows

Pritpal Matharu ●* and Bartosz Protas ●†

Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario L8S 4K1, Canada

We consider the question of fundamental limitations on the performance of eddy-viscosity closure models for turbulent flows, focusing on the Leith model for two-dimensional large-eddy simulation. Optimal eddy viscosities depending on the magnitude of the vorticity gradient are determined subject to minimum assumptions by solving PDE-constrained optimization problems defined such that the corresponding optimal large-eddy simulation best matches the filtered direct numerical simulation. First, we consider pointwise match in the physical space, and the main finding is that with a fixed cutoff wave number k_c, the performance of the large-eddy simulation systematically improves as the regularization in the solution of the optimization problem is reduced, and this is achieved with the optimal eddy viscosities exhibiting increasingly irregular behavior with rapid oscillations. Since the optimal eddy viscosities do not converge to a well-defined limit as the regularization vanishes, we conclude that in this case the problem of finding an optimal eddy viscosity does not in fact have a solution and is thus ill-posed. We argue that this observation is consistent with the physical intuition concerning closure problems. The second problem we consider involves matching time-averaged vorticity spectra over small wave numbers. It is shown to be better behaved and to produce physically reasonable optimal eddy viscosities. We conclude that while better behaved and hence practically more useful eddy viscosities can be obtained with stronger regularization or by matching quantities defined in a statistical sense, the corresponding large-eddy simulations will not achieve their theoretical performance limits.

I. INTRODUCTION

The closure problem is arguably one of the most important outstanding open problems in turbulence research. It touches upon some of the key basic questions concerning turbulent flows and at the same time has far-reaching consequences for many applications, most importantly, for how we simulate turbulent flows in numerous geophysical, biological, and engineering settings. Given the extreme spatiotemporal complexity of turbulent flows, accurate numerical solutions of the Navier-Stokes system even at modest Reynolds numbers require resolutions exceeding the capability of commonly accessible computational resources. To get around this difficulty, one usually relies on various simplified versions of the Navier-Stokes system obtained through different forms of averaging and/or filtering, such as the Reynolds-averaged Navier-Stokes (RANS) system and the large-eddy simulation (LES). However, such formulations are not closed, because these systems involve nonlinear terms representing the effect of unresolved subgrid stresses on the resolved variables. The "closure problem" thus consists in expressing these quantities in terms of resolved variables such that the RANS or LES system is closed.

---

*mathap1@mcmaster.ca; https://ms.mcmaster.ca/~mathap1/
†bprotas@mcmaster.ca; https://ms.mcmaster.ca/~bprotas/

In general, closure models in fluid mechanics are of two main types: algebraic, where there is an algebraic relationship expressing the subgrid stresses in terms of the resolved quantities, and differential, where this relationship involves an additional partial-differential equation (PDE) which needs to be solved together with the RANS or LES system. Most classical models are usually formulated based on some *ad hoc*, albeit well-justified, physical assumptions. There exists a vast body of literature concerning the design, calibration, and performance of such models in various settings. Since it is impossible to offer an even cursory survey of these studies here, we refer the reader to the well-known monographs [1–3] for an overview of the subject. Recently, there has been significant activity centered on learning new empirical closure models from data using methods of machine learning [4–9]. It is, however, fair to say that the field of turbulence modeling has been largely dominated by empiricism, and there is a consensus that the potential and limitations of even the most common models are still not well understood. Our study tackles this fundamental question, more specifically, how well certain common closure models can in principle perform if they are calibrated in a optimal way. We will look for an optimal, in a mathematically precise sense, form of a certain closure model and will conclude that, somewhat surprisingly, it does not in fact exist.

On the other hand, from the physical point of view, turbulence closure models are not meant to capture nonlinear transfer processes with pointwise accuracy, but rather to represent them in a certain average sense. The ill-posedness of the problem of optimally calibrating a closure model signaled above can thus be viewed as a consequence of the inability of the closure model to match the original solution pointwise in space and in time. More precisely, the optimal eddy viscosity exhibits unphysical high-frequency oscillations. In the present study we will use a mathematically systematic approach to illustrate this physical intuition and demonstrate how the ill-posedness arises. We will also show that the model calibration problem is in fact well behaved when the LES with a closure model is required to match quantities defined in the statistical rather than pointwise sense.

We are going to focus on an example from a class of widely used algebraic closure models, namely, the Smagorinsky-type eddy-viscosity models [10] for LES. More specifically, we will consider the Leith model [11–13] for two-dimensional (2D) turbulent flows. Like all eddy-viscosity closure models, the Leith model depends on one key parameter, which is the eddy viscosity typically taken to be a function of some flow variable. Needless to say, performance of such models critically depends on the form of this function. One specific question we are interested in is how accurately the LES equipped with such an eddy-viscosity closure model can at best reproduce solutions of the Navier-Stokes system obtained via direct numerical simulation (DNS). Another related question we will consider concerns reproducing certain statistical properties of Navier-Stokes flows in LES. We will address these questions by formulating them as PDE-constrained optimization problems where we will seek an optimal functional dependence of the eddy viscosity on the state variable. In the first problem we will require the corresponding LES to match the filtered DNS pointwise in space over a time window of several eddy turnover times, whereas in the second problem the LES will be required to match the time-averaged enstrophy spectrum of the Navier-Stokes flow for small wave numbers. By framing these questions in terms of optimization problems we will be able to find the best (in a mathematically precise sense) eddy viscosities, and this will in turn allow us to establish ultimate performance limitations for this class of closure models. We emphasize that the uniqueness of our approach is that by finding an optimal functional form of the eddy viscosity we identify, subject to minimum assumptions, an optimal structure of the nonlinearity in the closure model, which is fundamentally different, and arguably more involved, than calibrating one or more constants in a selected ansatz for the eddy viscosity. This formulation is also more general than common dynamic closure models and some formulations employing machine learning to deduce information about local properties of closure models from the DNS (see, e.g., [14]). Our goal is to understand what form the eddy viscosity needs to take in order to maximize the performance of the closure model in achieving a prescribed objective. The emphasis will be on methodology rather than on specific contributions to subgrid modeling.

The optimization problem in question has a nonstandard structure, but an elegant solution can be obtained using a generalization of the adjoint-based approach developed in Refs. [15,16]. In being based on methods of the calculus of variations, this approach thus offers a mathematically rigorous alternative to machine-learning methods which have recently become popular [4–9]. As a proof of the concept applicable to the problem considered here, this approach was recently adapted to find optimal closures in a simple one-dimensional (1D) model problem by Matharu and Protas [17]. Importantly, this approach involves a regularization parameter controlling the "smoothness" of the obtained eddy viscosity.

In the first problem, which involves matching the filtered DNS solution in the pointwise sense, we find optimal eddy viscosities for the Leith closure model in the LES systems with different filter cutoff wave numbers k_c. As this wave number increases and the filter width vanishes, the optimal eddy viscosity is close to zero, and the match between the predictions of the LES and the filtered DNS is nearly exact, as expected. On the other hand, for smaller cutoff wave numbers k_c the optimal eddy viscosity becomes highly irregular, whereas the match between the LES and DNS deteriorates, although it still remains much better than the match involving the LES with the standard Leith model or with no closure model at all. Interestingly, the optimal eddy viscosity reveals highly oscillatory behavior with alternating positive and negative values as the state variable increases. When the regularization in the solution of the optimization problems is reduced and the numerical resolution is refined at a fixed cutoff wave number, the frequency and amplitude of these oscillations are amplified, which results in an improved match against the DNS. Thus, in this limit the optimal eddy viscosity becomes increasingly oscillatory as a function of the state variable, which suggests that in the absence of regularization the problem of finding an optimal eddy viscosity does not in fact have a solution as the limiting eddy viscosity is not well defined. On the other hand, an arbitrarily regular eddy viscosity can be found when sufficient regularization is used in the solution of the optimization problem, but at the price of reducing the match against the DNS. While such smooth eddy viscosities may be more useful in practice, the corresponding LES models will not achieve their theoretical performance limits. In addition to this observation, our results also demonstrate how the best accuracy achievable by the LES with the considered closure model depends on the cutoff wave number of the filter, which sheds light on the fundamental performance limitations inherent in this closure model.

In our second problem, which involves matching the time-averaged vorticity spectrum of the filtered DNS, the obtained optimal eddy viscosity is more regular, and its key features remain essentially unchanged as the regularization in the solution of the optimization problems is reduced and the numerical resolution is refined. This demonstrates that the problem of optimally calibrating the closure model is better behaved when a suitable statistical quantity is used as the target. This is not surprising as such a formulation is in fact closer to the spirit of turbulence modeling.

The structure of the paper is as follows: in the next section we formulate our LES model and state the optimization problem defining the optimal eddy viscosity; in Sec. III we introduce an adjoint-based approach to the solution of the optimization problem and in Sec. IV discuss computational details; our results are presented in Sec. V whereas final conclusions are deferred to Sec. VI; some additional technical material is provided in the Appendix.

II. LARGE-EDDY SIMULATION AND OPTIMAL EDDY VISCOSITY

We consider 2D flows of viscous incompressible fluids on a periodic domain $\Omega := [0, 2\pi]^2$ over the time interval $[0, T]$ for some $T > 0$ (":=" means "equal to by definition"). Assuming the fluid is of uniform unit density $\rho = 1$, its motion is governed by the Navier-Stokes system written here in the vorticity form

$$\partial_t w + \nabla^\perp \psi \cdot \nabla w = \nu_N \Delta w - \alpha w + f_\omega \quad \text{in} \quad (0, T] \times \Omega, \tag{1a}$$

$$\Delta \psi = -w \quad \text{in} \quad (0, T] \times \Omega, \tag{1b}$$

$$w(t = 0) = w_0 \quad \text{in} \quad \Omega, \tag{1c}$$

where $w = -\nabla^\perp \cdot \boldsymbol{u}$, with $\nabla^\perp = [\partial_{x_2}, -\partial_{x_1}]^T$ and $\boldsymbol{u}$ the velocity field, is the vorticity component perpendicular to the plane of motion, ψ is the streamfunction, ν_N is the coefficient of the kinematic viscosity (for simplicity, we reserve the symbol ν for the eddy viscosity), and w_0 is the initial condition. System (1) is subject to two forcing mechanisms: a time-independent forcing f_ω which ensures that the flow remains in a statistical equilibrium and the Ekman friction $-\alpha w$ describing large-scale dissipation due to, for example, interactions with boundary layers arising in geophysical fluid phenomena. The forcing term is defined to act on Fourier components of the solution with wave numbers in the range $[k_a, k_b]$ for some $0 < k_a < k_b < \infty$, i.e.,

$$[\widehat{f_\omega}]_{\boldsymbol{k}} := \begin{cases} F, & k_a \leqslant |\boldsymbol{k}| \leqslant k_b, \\ 0, & \text{otherwise,} \end{cases} \tag{2}$$

where $[\widehat{f_\omega}]_{\boldsymbol{k}}$ is the Fourier component of f_ω with the wave vector $\boldsymbol{k}$ (hereafter hats "$\widehat{\cdot}$" will denote Fourier coefficients) and $F > 0$ is a constant parameter.

The phenomenology of 2D forced turbulence is described by the Kraichnan-Batchelor-Leith theory [11,18,19], which makes predictions about various physical characteristics of such flows. Their prominent feature, distinct from turbulent flows in three dimensions (3D), is the presence of a forward enstrophy cascade and an inverse energy cascade [20–24]. Here we will chose k_a and k_b such that the forcing term (2) will act on a narrow band of Fourier coefficients to produce a well-developed enstrophy cascade towards large wave numbers and a rudimentary energy cascade towards small wave numbers. The parameters ν_N, α, and F will be adjusted to yield a statistically steady state with enstrophy $\mathcal{E}(t) := \int_\Omega w^2(t, \mathbf{x})\, d\Omega$ fluctuating around a well-defined mean value $\mathcal{E}_0$. The initial condition ω_0 in (1c) will be chosen such that the evolution begins already in this statistically steady state at time $t = 0$.

A. The Leith closure model

The LES is obtained by applying a suitable low-pass filter G_δ, where $\delta > 0$ is its width, to the Navier-Stokes system (1) and defining the filtered variables $\widetilde{w} = G_\delta * w$ and $\widetilde{\psi} = G_\delta * \psi$ (" $*$ " denotes the convolution operation and hereafter we will use tilde "$\widetilde{\cdot}$" to represent filtered variables). For simplicity, we will employ a sharp low-pass spectral filter defined in terms of its Fourier-space representation as

$$[\widehat{G_\delta}]_{\boldsymbol{k}} := \begin{cases} 1, & |\boldsymbol{k}| \leqslant k_c, \\ 0, & \text{otherwise,} \end{cases} \tag{3}$$

where k_c is the largest resolved wave number such that the filter width is $\delta = 2\pi/k_c$. Since we normally have $k_b < k_c$, it follows that $\widetilde{f_\omega} = f_\omega$. Application of filter (3) to the vorticity equation (1a) yields $\partial_t \widetilde{w} + \widetilde{\nabla^\perp \psi \cdot \nabla w} = \nu_N \Delta \widetilde{w} - \alpha \widetilde{w} + f_\omega + M$, where the term M represents the effect of the unresolved subgrid quantities

$$M = \nabla^\perp \widetilde{\psi} \cdot \nabla \widetilde{w} - \widetilde{\nabla^\perp \psi \cdot \nabla w}. \tag{4}$$

Since expression (4) depends on the original unfiltered variables w and ψ, to close the filtered system the term M must be modeled in terms of an expression involving the filtered variables only. We will do this using the Leith model [11–13], which has a similar structure to the Smagorinsky model [10] widely used as a closure for 3D flows, but is derived considering the forward enstrophy cascade as the dominant mechanism in 2D turbulent flows. There is evidence for good performance of the Leith model in such flows [25,26]. Its preferred form is

$$M \approx \widetilde{M} = \nabla \cdot (\widetilde{\nu_L \nabla \widetilde{\omega}}), \tag{5}$$

in which $\widetilde{\omega}$ is the solution to the LES system [cf. (8)] and the eddy viscosity is assumed to be a linear function of the magnitude of the vorticity gradient, i.e.,

$$\nu_L(s) := (C_L\delta)^3 \sqrt{s} \quad \text{with} \quad s := |\nabla\widetilde{\omega}|^2 \in \mathcal{I} := [0, s_{\max}], \tag{6}$$

where the Leith constant $C_L = 1$ and $s_{\max} > 0$ is a sufficiently large number to be specified later. We will refer to $\mathcal{I}$ as the "state space" domain.

While in the original formulation of the Leith model the eddy viscosity is taken to be a linear function of $|\nabla\widetilde{\omega}|$ as in (6) [25,26], here we consider a general dependence of the eddy viscosity on $|\nabla\widetilde{\omega}|$ in the form

$$\nu(s) = [\nu_L(s) + \nu_0]\varphi\left(\frac{s}{s_{\max}}\right), \tag{7}$$

where $\nu_0 > 0$ and $\varphi : [0, 1] \to \mathbb{R}$ is a dimensionless function subject to some minimum only assumptions to be specified below. The parameter ν_0 is introduced to allow the eddy viscosity $\nu(s)$ to take nonzero values at $s = 0$, in contrast to Leith's original model (6). We remark that defining the eddy viscosity in terms of such a function φ ensures that ansatz (7) is dimensionally consistent. Making φ and ν functions of $|\nabla\widetilde{\omega}|^2$, rather than of $|\nabla\widetilde{\omega}|$, in (7) will simplify subsequent calculations. We add that ansatz (7) is used here to illustrate the approach, and in principle one could also consider other formulations parametrized by nondimensional functions. With the Leith model (5)–(7), the LES version of the 2D Navier-Stokes system (1) takes the form

$$\partial_t \tilde{w} + \widetilde{\nabla^\perp\widetilde{\psi} \cdot \nabla \tilde{w}} = \nabla \cdot ([\nu_N + \widetilde{\nu(s)}]\nabla\tilde{w}) - \alpha\bar{w} + f_w \quad \text{in} \quad (0, T] \times \Omega, \tag{8a}$$

$$\mathbf{\Delta}\widetilde{\psi} = -\widetilde{\omega} \qquad\qquad\qquad \text{in} \quad (0, T] \times \Omega, \tag{8b}$$

$$\widetilde{\omega}(t = 0) = \widetilde{\omega}_0 := \widetilde{w}_0 \qquad\qquad\qquad \text{in} \quad \Omega, \tag{8c}$$

where the initial condition is given as the filtered initial condition (1c) from the DNS system.

An equivalent form of equation (8a) can be obtained noting that with the form of the filter given in (3), the decomposition of the subgrid stresses (4) reduces to $M = \nabla^\perp\widetilde{\psi} \cdot \nabla \tilde{w} - \widetilde{\nabla^\perp\widetilde{\psi} \cdot \nabla w}$ [2]. As a result, the advection term in (8a) can be replaced with $\nabla^\perp\widetilde{\psi} \cdot \nabla\widetilde{\omega}$. While our numerical solution will be based on (8a), this second form will facilitate the derivations presented in Sec. III. We will assume that for all times $t \in [0, T]$ the filtered vorticity field $\widetilde{\omega}$ is in the Sobolev space $H_0^2(\Omega)$ of zero-mean functions with square-integrable second derivatives [27]. We stress the distinction between the fields $w, \tilde{w}, \widetilde{\omega}$, which represent, respectively, the solution of the DNS system (1), its filtered version, and the solution of the LES system (8).

B. Optimization formulation for eddy viscosity

We consider two formulations with the DNS field matched pointwise in space and in time, and in a certain statistical sense. First, the optimal eddy viscosity will be found as a minimizer of an error functional representing the mean-square error between observations of the filtered DNS, i.e., of the filtered solution $\tilde{w}(t, \mathbf{x})$ of the Navier-Stokes system (1), and observations the corresponding prediction $\widetilde{\omega}(t, \mathbf{x}; \varphi)$ of the LES model (8) with eddy viscosity ν. These observations are acquired at points $\mathbf{x}_i, i = 1, \ldots, M^2$, forming a uniform $M \times M$ grid in Ω with operators $H_i : H^2(\Omega) \longrightarrow \mathbb{R}$ defined as

$$(H_i\widetilde{\omega})(t) := \int_\Omega \delta(\mathbf{x} - \mathbf{x}_i)\widetilde{\omega}(t, \mathbf{x}) \, d\Omega = \widetilde{\omega}(t, \mathbf{x}_i), \quad i = 1, \ldots, M^2, \tag{9}$$

where $\delta(\cdot)$ is the Dirac delta distribution and observations $[H_i\widetilde{\omega}(\varphi)](t)$ of the LES solution are defined analogously (an integral representation of the observation operators will be convenient for the derivation of the solution approach for the optimization problem presented in Sec. III). The number of the observations points M^2 will be chosen such that $M \gtrsim k_c$, i.e., the observations will

resolve all flow features with wave numbers slightly higher than the cutoff wave number k_c in (3). The error functional then takes the form

$$\mathcal{J}_1(\varphi) := \frac{1}{2} \int_0^T \sum_{i=1}^{M^2} \{(H_i \widetilde{w})(t) - [H_i \widetilde{\omega}(\varphi)](t)\}^2 \, dt \tag{10}$$

and is understood as depending on the function φ parametrizing the eddy viscosity $\nu = \nu(s)$ via ansatz (7).

In the second formulation, the optimal eddy viscosity will be found by minimizing the error between the time-averaged vorticity spectra in the filtered DNS and predicted by the LES. For simplicity and with a slight abuse of notation, we will treat the wave number k as a continuous variable, i.e., we will assume that $\boldsymbol{k} \in \mathbb{R}^2$ rather than $\boldsymbol{k} \in \mathbb{Z}^2$; in the actual implementation one needs to account for the discrete nature of the wave vector $\boldsymbol{k}$. The vorticity spectrum predicted by the LES is then defined as

$$E_{\widetilde{\omega}}(t, k) := \frac{1}{2} \int_{\mathscr{C}(k)} |\widehat{\widetilde{\omega}}(t, \boldsymbol{k})|^2 \, dS(\boldsymbol{k}), \quad \forall t, k \geqslant 0, \tag{11}$$

where $\widehat{\widetilde{\omega}}(t, \boldsymbol{k})$ is the Fourier transform of $\widetilde{\omega}(t, \mathbf{x})$ and $\mathscr{C}(k) := \{\boldsymbol{k} \in \mathbb{R}^2 : |\boldsymbol{k}| = k\}$ a circle with radius k in the 2D plane. The vorticity spectrum $E_w(t, k)$ in the (filtered) DNS is defined analogously. Denoting $[f]_T := (1/T) \int_0^T f(t) \, dt$ the time average of a function $f : [0, T] \to \mathbb{R}$, the error functional is defined as

$$\mathcal{J}_2(\varphi) := \frac{1}{4} \int_{k=0}^{k_c} \{[E_{\widetilde{\omega}}(\cdot, k; \varphi)]_T - [E_w(\cdot, k)]_T\}^2 \, dk, \tag{12}$$

with matching performed up to the cutoff wave number k_c.

The form of equation (8a) suggests that $\nu = \nu(s)$, and hence also $\varphi = \varphi(s/s_{\max})$, must be at least piecewise C^1 functions on $\mathcal{I}$ and [0,1], respectively. However, as will become evident in Sec. III, our solution approach imposes some additional regularity requirements, namely, $\nu = \nu(s)$ needs to be piecewise C^2 on $\mathcal{I}$ with the first and third derivatives vanishing at $s = 0$, $s_{\max}$. Since gradient-based solution approaches to PDE-constrained optimization problems are preferably formulated in Hilbert spaces [28], we shall look for an optimal function φ parametrizing the optimal eddy viscosity as an element of the following linear space which is a subspace of the Sobolev space $H^2(\mathcal{I})$:

$$\mathcal{S} := \left\{ \varphi \in C^3([0, 1]) : \frac{d}{d\xi} \varphi(\xi) = \frac{d^3}{d\xi^3} \varphi(\xi) = 0 \text{ at } \xi = 0, 1 \right\}. \tag{13}$$

Then the problem of finding an optimal eddy viscosity in the two formulations becomes

$$\breve{\varphi} := \arg \min_{\varphi \in \mathcal{S}} \mathcal{J}_j(\varphi), \quad j = 1, 2, \tag{14}$$

where the optimal eddy viscosity $\breve{\nu}$ is deduced from $\breve{\varphi}$ via ansatz (7). Our approach to solving this problem is outlined in the next section.

III. ADJOINT-BASED OPTIMIZATION

We focus here on solution of the optimization problem in the first formulation, i.e., for $j = 1$ in (14), with the error functional given in (10). Essentially the same approach can also be used to solve the second optimization problem with the error functional (12) and required modifications are discussed in the Appendix. We formulate our approach in the continuous ("optimize-then-discretize") setting [29] and adopt the strategy developed and validated by Matharu and Protas [17]. Here we only summarize its key steps and refer the reader to that study for further details. A local solution of problem (10), (13), and (14) can be found using an iterative gradient-based minimization

approach as $\breve{\varphi} = \lim_{n \to \infty} \varphi^{(n)}$, where

$$\varphi^{(n+1)} = \varphi^{(n)} - \tau^{(n)} \nabla_\varphi \mathcal{J}_1(\varphi^{(n)}), \quad n = 0, 1, \ldots,$$

$$\varphi^{(0)} = \varphi_0, \tag{15}$$

in which $\varphi^{(n)}$ is the approximation of the optimal function $\breve{\varphi}$ at the nth iteration (which can be used to construct the corresponding approximation $\nu^{(n)}$ of the optimal eddy viscosity), $\nabla_\varphi \mathcal{J}_1(\varphi)$ is the gradient of the error functional (10) with respect to φ, $\tau^{(n)}$ is the step length along the descent direction, and φ_0 is an initial guess usually suggested by some form of the eddy viscosity.

A central element of algorithm (15) is the gradient $\nabla_\varphi \mathcal{J}_1(\varphi)$. In many problems of PDE-constrained optimization it can be conveniently expressed using solutions of suitably defined adjoint equations [29]. However, the present optimization problem (10), (13), and (14) has a nonstandard structure because the control variable $\varphi(s/s_{\max})$ is a function of the *dependent* variable $s = |\nabla\widetilde{\omega}|^2$ in system (8). On the other hand, in its standard formulation adjoint analysis allows one to obtain expressions for gradients depending on the *independent* variables in the problem (here t and $\mathbf{x}$). This difficulty was overcome in Refs. [15,16], which generalized adjoint analysis of PDE systems to problems of the type (10), (13), and (14) by introducing a suitable change of variables. For convenience we will denote $\sigma := s/s_{\max}$.

The Gâteaux (directional) differential of the error functional (10) with respect to φ, defined by $\mathcal{J}_1'(\varphi; \varphi') := \lim_{\epsilon \to 0} \epsilon^{-1}[\mathcal{J}_1(\varphi + \epsilon\varphi') - \mathcal{J}_1(\varphi)]$, is defined as

$$\mathcal{J}_1'(\varphi; \varphi') = \int_0^T \int_\Omega \sum_{i=1}^{M^2} H_i^* \{[H_i \widetilde{\omega}(\varphi)](t) - (H_i \widetilde{w})(t)\} \widetilde{\omega}'(t, \mathbf{x}; \varphi, \varphi') \, d\mathbf{x} \, dt, \tag{16}$$

where $\varphi' \in \mathcal{S}$ is an arbitrary perturbation of the control variable φ, $\widetilde{\omega}'(t, \mathbf{x}; \varphi, \varphi')$ satisfies the system

$$\mathcal{K} \begin{bmatrix} \widetilde{\omega}' \\ \widetilde{\psi}' \end{bmatrix} := \begin{bmatrix} \partial_t \widetilde{\omega}' + \nabla^\perp \widetilde{\psi}' \cdot \nabla\widetilde{\omega} + \nabla^\perp \widetilde{\psi} \cdot \nabla\widetilde{\omega}' + \alpha\widetilde{\omega}' \\ -\nabla \cdot \left[2(\nabla\widetilde{\omega} \cdot \nabla\widetilde{\omega}') \left(\frac{dv}{ds} \varphi \nabla\widetilde{\omega} + \frac{\nu_L + \nu_0}{s_{\max}} \frac{d\varphi}{d\sigma} \nabla\widetilde{\omega} \right) + (\nu_N + \nu)\nabla\widetilde{\omega}' \right] \\ \Delta\widetilde{\psi}' + \widetilde{\omega}' \end{bmatrix}$$

$$= \begin{bmatrix} \nabla \cdot [(\nu_L + \nu_0)\varphi' \nabla\widetilde{\omega}] \\ \\ 0 \end{bmatrix}, \tag{17a}$$

$$\widetilde{\omega}'(t = 0, \mathbf{x}) = 0, \tag{17b}$$

obtained as linearization of the LES system (8) and $H_i^* : \mathbb{R} \longrightarrow H^{-2}(\Omega)$, $i = 1, \ldots, M^2$, are the adjoints of the observation operators H_i [cf. Eq. (9)], given by

$$\forall \xi \in \mathbb{R}, \quad (H_i^* \xi) := \delta(\mathbf{x} - \mathbf{x}_i)\xi, \quad i = 1, \ldots, M^2. \tag{18}$$

In order to extract the gradient $\nabla_\varphi \mathcal{J}_1$ from the Gâteaux differential (16), we note that this derivative is a bounded linear functional when viewed as a function of φ' and invoke the Riesz representation theorem [30] to obtain

$$\mathcal{J}_1'(\varphi; \varphi') = \left\langle \nabla_\varphi^{H^2} \mathcal{J}_1, \varphi' \right\rangle_{H^2([0,1))} = \left\langle \nabla_\varphi^{L^2} \mathcal{J}_1, \varphi' \right\rangle_{L^2([0,1))}, \tag{19}$$

where the inner product in the space $H^2([0, 1))$ is defined as

$$\langle p_1, p_2 \rangle_{H^2([0,1))} = \int_0^1 p_1 \, p_2 + \ell_1^2 \frac{dp_1}{d\sigma} \frac{dp_2}{d\sigma} + \ell_2^4 \frac{d^2 p_1}{d\sigma^2} \frac{d^2 p_2}{d\sigma^2} \, d\sigma, \tag{20}$$

in which ℓ_1 and ℓ_2 are length-scale parameters. While for all values of $\ell_1, \ell_2 \in (0, \infty)$ the inner products (20) are equivalent (in the sense of norm equivalence), these two parameters play a very important role in regularization of solutions to the optimization problem (10)–(14). In (15) we

require the gradient in the space $H^2([0, 1])$, i.e., $\nabla_\varphi \mathcal{J}_1 = \nabla_\varphi^{H^2} \mathcal{J}_1$, but it is convenient to first derive the gradient with respect to the L^2 topology.

Introducing *adjoint fields* $\widetilde{\omega}^*$ and $\widetilde{\psi}^*$, we can define the following duality-pairing relation:

$$\left(\mathcal{K} \begin{bmatrix} \widetilde{\omega}' \\ \widetilde{\psi}' \end{bmatrix}, \begin{bmatrix} \widetilde{\omega}^* \\ \psi^* \end{bmatrix} \right) := \int_0^T \int_\Omega \mathcal{K} \begin{bmatrix} \widetilde{\omega}' \\ \widetilde{\psi}' \end{bmatrix} \cdot \begin{bmatrix} \widetilde{\omega}^* \\ \widetilde{\psi}^* \end{bmatrix} d\mathbf{x}\, dt =$$

$$\int_0^T \int_\Omega \begin{bmatrix} \widetilde{\omega}' \\ \widetilde{\psi}' \end{bmatrix} \cdot \mathcal{K}^* \begin{bmatrix} \widetilde{\omega}^* \\ \widetilde{\psi}^* \end{bmatrix} d\mathbf{x}\, dt = \left(\begin{bmatrix} \widetilde{\omega}' \\ \widetilde{\psi}' \end{bmatrix}, \mathcal{K}^* \begin{bmatrix} \widetilde{\omega}^* \\ \psi^* \end{bmatrix} \right), \tag{21}$$

where integration by parts was performed with respect to both space and time [noting the periodic boundary conditions and the initial condition (17b)] and the *adjoint system* has the form

$$\mathcal{K}^* \begin{bmatrix} \widetilde{\omega}^* \\ \widetilde{\psi}^* \end{bmatrix} := \begin{bmatrix} -\partial_t \widetilde{\omega}^* - \boldsymbol{\nabla}^\perp \widetilde{\psi} \cdot \boldsymbol{\nabla} \widetilde{\omega}^* + \alpha \widetilde{\omega}^* + \widetilde{\psi}^* \\ -\boldsymbol{\nabla} \cdot \left[2 \left(\boldsymbol{\nabla} \widetilde{\omega} \cdot \boldsymbol{\nabla} \widetilde{\omega}^* \right) \left(\frac{dv}{ds} \varphi \boldsymbol{\nabla} \widetilde{\omega} + \frac{v_L + v_0}{s_{max}} \frac{d\varphi}{d\sigma} \boldsymbol{\nabla} \widetilde{\omega} \right) + (v_N + v) \boldsymbol{\nabla} \widetilde{\omega}^* \right] \\ \boldsymbol{\Delta} \widetilde{\psi}^* - \boldsymbol{\nabla}^\perp \cdot (\widetilde{\omega}^* \boldsymbol{\nabla} \widetilde{\omega}) \end{bmatrix} = \begin{bmatrix} W \\ 0 \end{bmatrix}, \tag{22a}$$

$$\widetilde{\omega}^*(t = T, \mathbf{x}) = 0, \tag{22b}$$

with the source term $W(t, \mathbf{x}) := \sum_{i=1}^{M^2} H_i^* \{[H_i \widetilde{\omega}(\varphi)](t) - (H_i \widetilde{w})(t)\}$. Combining (17), (21), and (22), we then arrive at

$$\left(\begin{bmatrix} \widetilde{\omega}' \\ \widetilde{\psi}' \end{bmatrix}, \mathcal{K}^* \begin{bmatrix} \widetilde{\omega}^* \\ \widetilde{\psi}^* \end{bmatrix} \right) = \overbrace{\int_0^T \int_\Omega W(t, \mathbf{x})\, \widetilde{\omega}'\, d\mathbf{x}\, dt}^{\mathcal{J}_1{}'(\varphi; \varphi')}$$

$$= -\int_0^T \int_\Omega (v_L + v_0)\, (\boldsymbol{\nabla} \widetilde{\omega} \cdot \boldsymbol{\nabla} \widetilde{\omega}^*)\, \varphi'\, d\mathbf{x}\, dt, \tag{23}$$

from which we obtain an expression for the Gâteaux differential

$$\mathcal{J}_1{}'(\varphi; \varphi') = -\int_0^T \int_\Omega (v_L + v_0)\, (\boldsymbol{\nabla} \widetilde{\omega} \cdot \boldsymbol{\nabla} \widetilde{\omega}^*)\, \varphi'\, d\mathbf{x}\, dt,$$

with the perturbation φ' now appearing explicitly as a factor. However, this expression is still not consistent with the Riesz form (19), which requires integration with respect to s over [0,1]. In order to perform the required change of variables, we make the substitution $\varphi'(\boldsymbol{\nabla} \widetilde{\omega} \cdot \boldsymbol{\nabla} \widetilde{\omega}) = \int_0^1 \delta(\frac{\boldsymbol{\nabla} \widetilde{\omega} \cdot \boldsymbol{\nabla} \widetilde{\omega}}{s_{max}} - \sigma)\, \varphi'(\sigma)\, d\sigma$. Fubini's theorem then allows us to swap the order of integration such that the Gâteaux differential (16) is finally recast in the Riesz form (19) as an integral with respect to σ

$$\mathcal{J}_1{}'(\varphi; \varphi') = \int_0^1 \left[-\int_0^T \int_\Omega \delta\left(\frac{\boldsymbol{\nabla} \widetilde{\omega} \cdot \boldsymbol{\nabla} \widetilde{\omega}}{s_{max}} - \sigma \right) (v_L + v_0)\, \boldsymbol{\nabla} \widetilde{\omega} \cdot \boldsymbol{\nabla} \widetilde{\omega}^*\, d\mathbf{x}\, dt \right] \varphi'(\sigma)\, d\sigma. \tag{24}$$

The gradient defined with respect to the L^2 topology is then deduced from this expression as

$$\nabla_\varphi^{L^2} \mathcal{J}_1(\sigma) = -\int_0^T \int_\Omega \delta\left(\frac{\boldsymbol{\nabla} \widetilde{\omega} \cdot \boldsymbol{\nabla} \widetilde{\omega}}{s_{max}} - \sigma \right) (v_L + v_0)\, \boldsymbol{\nabla} \widetilde{\omega} \cdot \boldsymbol{\nabla} \widetilde{\omega}^*\, d\mathbf{x}\, dt. \tag{25}$$

The L^2 gradient given in (25) may in principle be discontinuous as a function of s and hence will not ensure the regularity required of the optimal eddy viscosity; cf. Sec. II B. To circumvent this problem, we define a Sobolev gradient using the Riesz relations (19) to identify the H^2 inner product (20) with expression (24) for the Gâteaux differential. Integrating by parts with respect to

σ and noting that the perturbation $\varphi' \in \mathcal{S}$ is arbitrary, we obtain the Sobolev gradient $\nabla_\varphi^{H^2} \mathcal{J}$ as a solution of the elliptic boundary-value problem

$$\left[\text{Id} - \ell_1^2 \frac{d^2}{d\sigma^2} + \ell_2^4 \frac{d^4}{d\sigma^4} \right] \nabla_\varphi^{H^2} \mathcal{J}_1(\sigma) = \nabla_\varphi^{L^2} \mathcal{J}_1(\sigma), \quad \sigma \in [0, 1], \tag{26a}$$

$$\left. \frac{d^{(1)} \left(\nabla_\varphi^{H^2} \mathcal{J}_1 \right)}{d\sigma^{(1)}} \right|_{\sigma=0,1} = \left. \frac{d^{(3)} \left(\nabla_\varphi^{H^2} \mathcal{J}_1 \right)}{d\sigma^{(3)}} \right|_{\sigma=0,1} = 0. \tag{26b}$$

The choice of the boundary conditions in (26b) ensures the vanishing of all the boundary terms resulting from the integration by parts. There is in fact some freedom in how to cancel these terms, and the choice in (26b) is arguably the least restrictive. As argued in Sec. II A, we allow the eddy viscosity $\nu(s)$ to take nonzero values at $s = 0$, so the corresponding Sobolev gradient should not vanish at $\sigma = 0$ such that it can modify the value of $\varphi(0)$, which turns out to be important in practice; cf. Sec. V. Thus, the choice of boundary conditions at $\sigma = 0$ provided in (26b) is necessary. On the other hand, the choice of the boundary conditions at $\sigma = 1$ has been found to have little effect on the gradient and on the obtained results provided $s_{\max}$ is sufficiently large. Therefore, the form of these boundary conditions given in (26b) is justified by simplicity. The boundary conditions (26b) are the reason for the presence of additional constraints in the definition of space $\mathcal{S}$ in (13).

Determination of the Sobolev gradients $\nabla_\varphi^{H^2} \mathcal{J}_1$ based on the L^2 gradients $\nabla_\varphi^{L^2} \mathcal{J}_1$ by solving system (26) can be viewed as low-pass filtering of the latter gradient using a nonsharp filter (as discussed by Protas *et al.* [28], this can be seen representing the operator $[\text{Id} - \ell_1^2 (d^2/d\sigma^2) + \ell_2^4(d^4/d\sigma^4)]^{-1}$ in the Fourier space). The parameters ℓ_1 and ℓ_2 serve as cutoff length scales representing the wavelengths of the finest features retained in the gradients $\nabla_\varphi^{H^2} \mathcal{J}_1$ such that increasing ℓ_1 and ℓ_2 has the effect of making the Sobolev gradient "smoother" and vice versa. Thus, ℓ_1 and ℓ_2 are "knobs" which can be tuned to control the regularity of the optimal eddy viscosities obtained as solutions of the problem (10)–(14).

Since by construction $\nabla_\varphi^{H^2} \mathcal{J}_1 \in \mathcal{S}$, choosing the initial guess in (15) such that $\varphi_0 \in \mathcal{S}$ will ensure that $\varphi^{(0)}, \varphi^{(1)}, \dots, \check{\varphi} \in \mathcal{S}$. At each step in (15) an optimal step size $\tau^{(n)}$ can be found by solving the following line-minimization problem [31]

$$\tau^{(n)} = \underset{\tau > 0}{\arg \min} \, \mathcal{J}_1[\varphi^{(n)} - \tau \, \nabla_\varphi \mathcal{J}_1(\varphi^{(n)})]. \tag{27}$$

Numerical implementation of the approach outlined above is discussed in the next section.

IV. COMPUTATIONAL APPROACH

The evaluation of the Sobolev gradient $\nabla_\varphi^{H^2} \mathcal{J}_1$ requires the numerical solutions of the LES system (8) and the adjoint system (22) followed by the solution of problem (26). For the first two systems we use a standard Fourier pseudospectral method in combination with a CN/RKW3 time-stepping technique introduced by Le and Moin [32], which give spectrally accurate results in space and a globally second-order accuracy in time. The spatial domain is discretized using $N_x = 256$ equispaced grid points in each direction. Since the eddy viscosity $\nu = \nu(s)$ and the function $\varphi(s/s_{\max})$ are state-dependent, we also need to discretize the state domain $\mathcal{I}$ [cf. (7)], which is done using N_s Chebyshev points (values of N_s are provided in Table I). We use Chebyshev differentiation matrices to perform differentiation with respect to s and the eddy viscosity $\nu(s)$, and its derivatives are interpolated from state space $\mathcal{I}$ to the spatial domain Ω using the barycentric formulas [33]. The boundary-value problem (26) is solved using a method based on ultraspherical polynomials available in the `chebop` feature of Chebfun [34]. Solution of the 2D Navier-Stokes system (1) is dealiased using Gaussian filtering based on the $3/2$ rule [35]; however, this is unnecessary for the LES system (8) due to the aggressive filtering applied. To ensure that aliasing errors resulting from the presence of state-dependent viscosity are eliminated, the adjoint system (22) is solved using twice as many grid points $2N_x$ in each direction.

TABLE I. Summary information about the different cases considered when solving optimization problem (14) with $j = 1$.

Case	k_c	N_s	ℓ_1	ℓ_2	φ_0	$\mathcal{J}_1(\varphi_0)$	$\mathcal{J}_1(\varphi^{(\infty)})$	r
A	30	64	10^4	10^3	No closure	4.398×10^{-7}	1.492×10^{-7}	8.999×10^{-8}
B	25	64	10^4	10^3	No closure	1.951×10^{-5}	2.450×10^{-6}	1.572×10^{-6}
C	20	64	10^4	10^3	No closure	3.635×10^{-4}	6.217×10^{-5}	4.468×10^{-5}
D	20	128	10^3	10^2	Case C	6.217×10^{-5}	2.001×10^{-5}	1.239×10^{-5}
E	20	256	10^1	10^0	Case D	2.333×10^{-5}	1.450×10^{-5}	8.723×10^{-6}

Evaluation of the L^2 gradient (25) requires nonstandard integration over level sets as described by Bukshtynov and Protas [16]. While for simplicity a simple gradient approach was presented in (15), in practice we use the Polak-Ribière variant of the conjugate-gradient method to accelerate convergence. For the line minimization problem (27), the standard Brent's algorithm is used [36]. The consistency and accuracy of the formulation and of the entire computational approach was validated using a standard suite of tests as was done by Matharu and Protas [17].

V. RESULTS

The results obtained by solving optimization problem (14) with error functionals (10) and (12) are presented in Secs. V A and V B. Our computations are based on a flow problem defined by the following parameters: $\nu_N = 1 \times 10^{-2}$, $\alpha = 1 \times 10^{-3}$, $F = 5$, and $k_a = k_b = 4$. In the first optimization problem we fix $M = 32$ in (10), which is slightly larger than the largest cutoff wave number k_c we consider (cf. Table I) and therefore ensures that the optimal eddy viscosity is determined based on all available flow information, and $T = 20 \approx 30 t_e$, where $t_e := [\int_0^T \mathcal{E}(t) \, dt / (8\pi^2 T)]^{-1/2}$ is the eddy turnover time [22]. We emphasize that the key insights provided by our computations do not depend on the particular choice of T, as long as it remains of comparable magnitude to the value given above.

A. Matching the DNS pointwise in space and time: Results for the optimization problem with error functional (10)

Our first set of results addresses the effect of the cutoff wave number k_c. They are obtained by solving problem (14) with $j = 1$ for decreasing values of $k_c = 30, 25, 20$ while retaining fixed values of the regularization parameters ℓ_1, ℓ_2 and a fixed resolution N_s in the state space $\mathcal{I}$; cf. cases A, B, and C in Table I. In each case the optimization problem is solved using the initial guess $\varphi_0(s/s_{\max}) \equiv 0$ corresponding to no closure model at all. The dependence of the error functional $\mathcal{J}_1(\nu^{(n)})$ on iterations n in the three cases is shown in Fig. 1(a), where we see that the mean-square errors between the DNS and the optimal LES increase as the cutoff wave number k_c is decreased, and the largest relative reduction of the error is achieved in case C with the smallest k_c. While minimization in problem (14) is performed with respect to the nondimensional function φ [cf. (7)], we focus here on the corresponding optimal eddy viscosities $\check{\nu} = \check{\nu}(s)$ shown in Fig. 1(b). Since small values of s are attained more frequently in the flow [cf. the probability density function (PDF) of $\sqrt{s}$ embedded in the figure], the horizontal axis is scaled as $\sqrt{s}$, which magnifies the region of small values of s. We see that for the largest cutoff wave number $k_c = 30$ the optimal eddy viscosity is close to zero over the entire range of s. However, for decreasing k_c the optimal eddy viscosity exhibits oscillations of increasing magnitude. We note that values of $s \gtrsim 50$ occur very rarely in the flow, and hence the gradient (25) provides little sensitivity information for s in this range. Thus, the behavior of $\check{\nu}(s)$ for $s \gtrsim 50$ is an artifact of the regularization procedure defined in (26) and is not physically relevant.
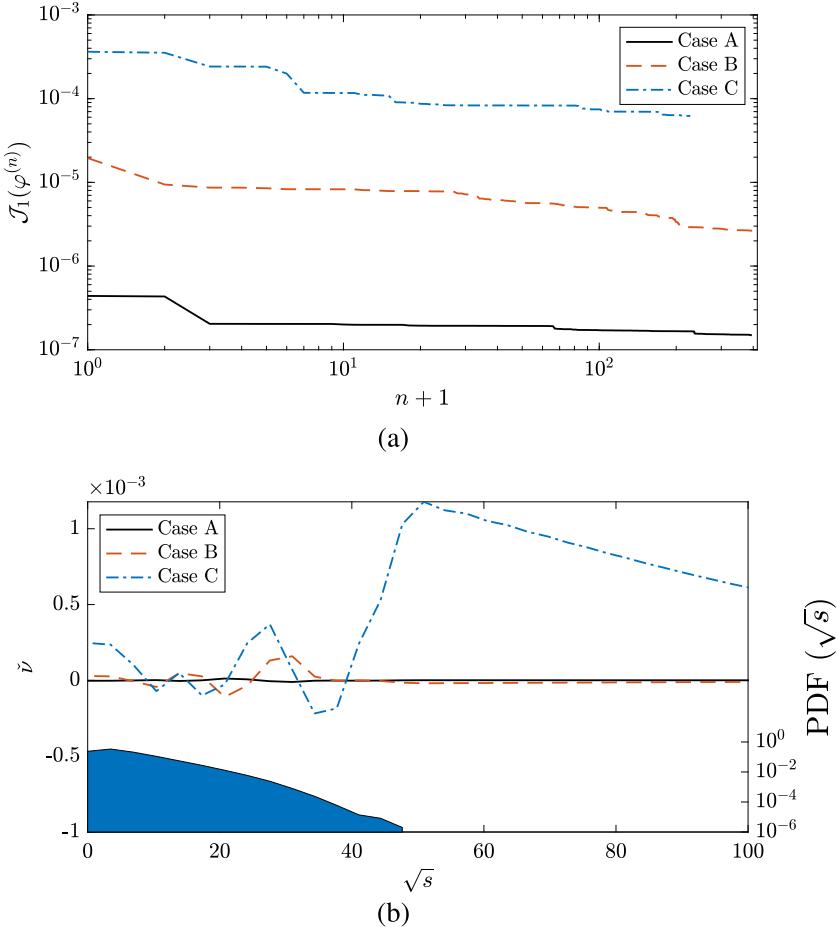
FIG. 1. (a) Dependence of the functional $\mathcal{J}_1(\varphi^{(n)})$ on the iteration n and (b) dependence of the corresponding optimal eddy viscosity $\breve{\nu}$ on $\sqrt{s}$ for cases A, B, and C; cf. Table I. Panel (b) also shows the PDF of $\sqrt{s}$ in case C.

In order to provide additional insights about the properties of the optimal eddy viscosity, our second set of results is obtained as solutions of problem (14) with $j = 1$ using a fixed $k_c = 20$ and progressively reduced regularization achieved by decreasing the parameters ℓ_1, ℓ_2 while simultaneously refining the resolution N_s in the state space $\mathcal{I}$; cf. cases C, D, and E in Table I. Optimization problems with weaker regularization are solved using the optimal eddy viscosity obtained with stronger regularization as the initial guess. From the normalized error functionals shown as functions of iterations in Fig. 2(a), we see that as regularization is reduced, the mean-square errors between the optimal LES and the DNS become smaller and approach a certain nonzero limit; cf. Table I. As is evident from Fig. 2(b), this is achieved with the corresponding optimal eddy viscosities developing oscillations with an ever increasing frequency. More precisely, each time the regularization parameters ℓ_1, ℓ_2 are reduced and the resolution N_s is refined, a new oscillation with a higher frequency appears in the optimal eddy viscosity $\breve{\nu}(s)$ (in fact, in each case, this is the highest-frequency oscillation which can be represented on a grid with N_s points).

In order to assess how well the solutions of the LES system (8) with the optimal eddy viscosities $\breve{\nu}$ shown in Figs. 1(b) and 2(b) approximate the solution of the Navier-Stokes system (1), in Figs. 3(a)
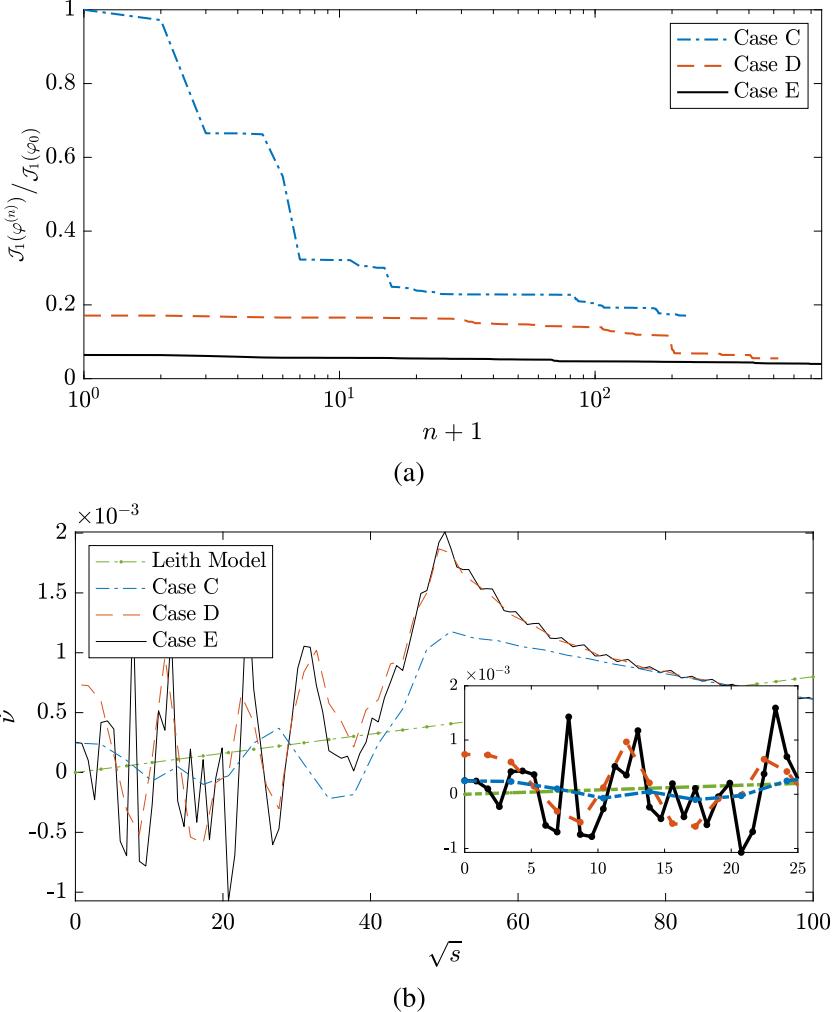
FIG. 2. (a) Dependence of the normalized functional $\mathcal{J}_1(\varphi^{(n)})/\mathcal{J}_1(\varphi_0)$, with $\mathcal{J}_1(\varphi_0)$ from case C, on the iteration n and (b) dependence of the corresponding optimal eddy viscosity $\check{\nu}$ on $\sqrt{s}$ for cases C, D, and E; cf. Table I. The inset in panel (b) shows magnification of the region $\sqrt{s} \in [0, 25]$. Panel (b) also shows the Leith model with $k_c = 20$ and the eddy viscosity $\nu_L(s)$; cf. (6).

and 3(b) we show the time evolution of the quantity $\log_{10}|1 - \mathcal{C}(t)|$ where

$$\mathcal{C}(t) := \frac{1}{||\widetilde{w}(t)||_{L^2(\Omega)} \, ||\widetilde{\omega}(t)||_{L^2(\Omega)}} \int_{\Omega} \widetilde{w}(t, \mathbf{x}) \, \widetilde{\omega}(t, \mathbf{x}) \, d\Omega \qquad (28)$$

is the normalized correlation between the two flows. For a more comprehensive assessment, these results are shown for $t \in [0, 2T]$, i.e., for times up to twice longer than the "training window" $[0, T]$ used in the optimization problem (14). In Fig. 3(b) we also present the results obtained for $k_c = 20$ with an optimal closure model based on the linear stochastic estimator introduced by Langford and Moser [37]. Since at early times correlation $\mathcal{C}(t)$ reveals exponential decay corresponding to the exponential divergence of the LES flow from the DNS, this effect can be quantified by approximating the correlation as $\mathcal{C}(t) \approx \bar{\mathcal{C}}(t) := \mathcal{C}_0 e^{-rt}$, where $\mathcal{C}_0 = 1$ follows the fact that $\widetilde{\omega}_0 \equiv \widetilde{w}_0$,
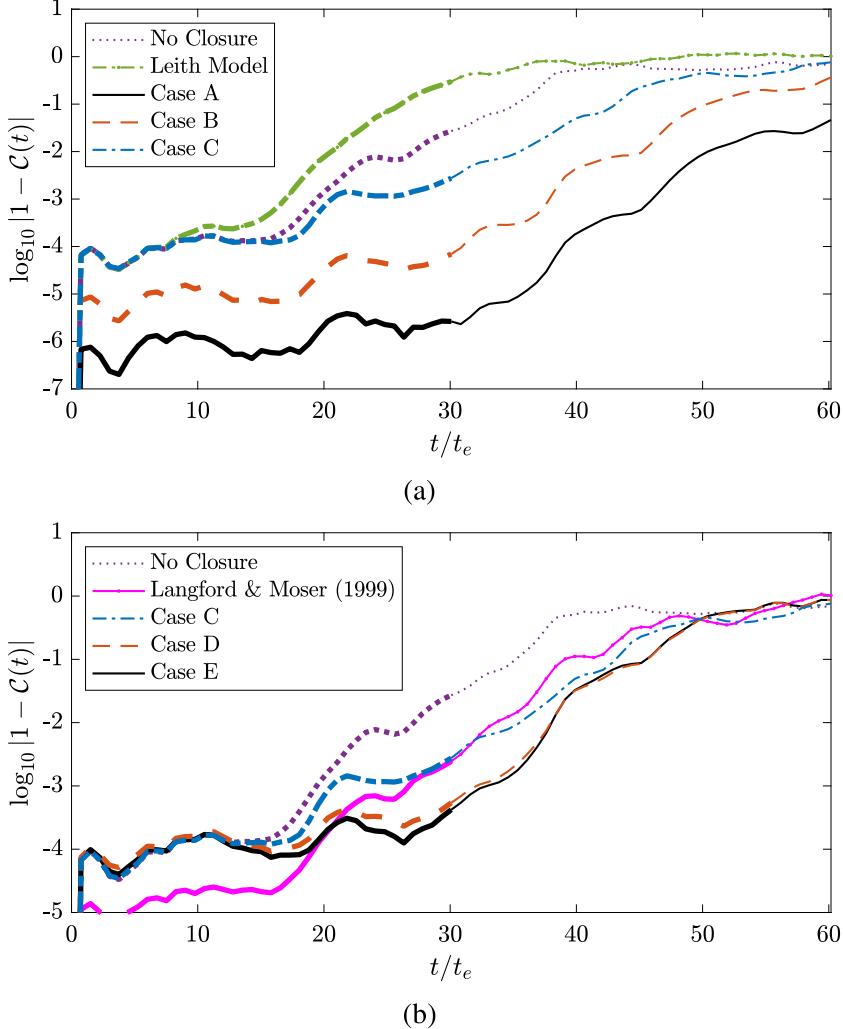
(a)

(b)

FIG. 3. Adjusted normalized correlations (28) for the LES with (a) no closure and the optimal eddy viscosity in cases A, B, and C, and (b) no closure and the optimal eddy viscosity in cases C, D, and E. The correlation is also shown for the Leith model with $k_c = 20$ and the eddy viscosity $\nu_L(s)$ [cf. (6)] in (a) and for an optimal closure model based on the stochastic estimator [37] in (b). Thick and thin lines correspond to, respectively, time in the "training window" ($t \in [0, T]$) and beyond this window ($t \in (T, 2T]$).

whereas the decay rate r is obtained from a least-squares fit over the time window $[0, T]$. The decay rates r obtained in this way are collected in Table I.

Finally, in order to provide insights about how the closure model with the optimal eddy viscosity acts in the physical space, in Figs. 4(a), 4(b), and 4(d) we show the vorticity field $\widetilde{\omega}(T, \mathbf{x})$, the corresponding state variable $s(T, \mathbf{x})$ [cf. (7)], and the spatial distribution $\check{\nu}(s(T, \mathbf{x}))$ of the optimal eddy viscosity obtained in case E; for comparison, the spatial distribution of the eddy viscosity $\nu_L(s(T, \mathbf{x}))$ in the Leith model [cf. (6) with $\delta = 0.02$] is shown in Fig. 4(c) (the fields are shown in the entire domain, i.e., for $\mathbf{x} \in \Omega$, at the end of the training window). We see that while the vorticity and state-variable fields vary smoothly, this is also the case for the spatial distribution of the eddy viscosity $\nu_L(s(T, \mathbf{x}))$ in the Leith model. On the other hand, the spatial distribution of the optimal eddy viscosity $\check{\nu}(s(T, \mathbf{x}))$ exhibits rapid variations, which is consistent with the results presented
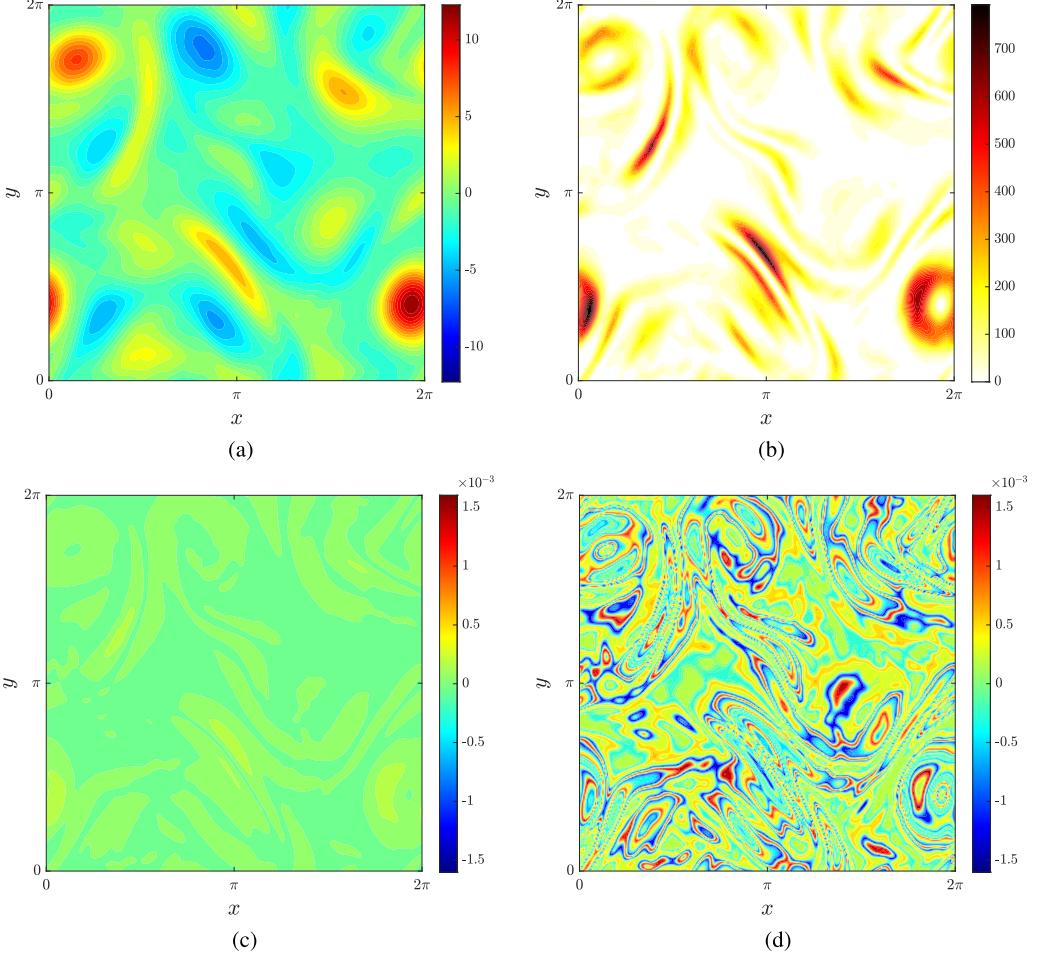
FIG. 4. For case E we show (a) the vorticity field $\widetilde{\omega}(T, \mathbf{x})$, $\mathbf{x} \in \Omega$, (b) the corresponding state variable $s(T, \mathbf{x})$ [cf. (7)], and the spatial distribution of (c) the eddy viscosity $\nu_L(s(T, \mathbf{x}))$ in the Leith model [cf. (6)] with $\delta = 0.02$, and (d) the optimal eddy viscosity $\breve{\nu}(s(T, \mathbf{x}))$ [cf. Fig. 2(b)], all shown at the end of the training window for $t = T$. For better comparison the same color scale is used in panels (c) and (d).

in Fig. 2(b). In particular, positive and negative values of $\breve{\nu}(s(T, \mathbf{x}))$, corresponding to localized dissipation and injection of enstrophy, tend to form concentric bands in some low-vorticity regions of the flow domain. The time evolution of the vorticity field in the DNS, LES with no closure model, and LES with the optimal eddy viscosity (case E) are available together with an animated version of Fig. 3(b) in the Supplemental Material as Movie 1 [38]. An animation representing the time evolution of the fields shown in Fig. 4 for $t \in [0, 2T]$ is available in the Supplemental Material as Movie 2 [38].

B. Matching the DNS in an average sense: Results for the optimization problem with error functional (12)

Now we review the results obtained by solving optimization problem (14) for $j = 2$ with a fixed cutoff wave number $k_c = 20$ and with two sets of parameters determining regularization (ℓ_1 and ℓ_2) and the resolution in the state space $\mathcal{I}$ (N_s); cf. cases F and G in Table II. We remark that the

TABLE II. Summary information about the different cases considered when solving optimization problem (14) with $j = 2$.

Case	k_c	N_s	ℓ_1	ℓ_2	φ_0	$\mathcal{J}_2(\varphi_0)$	$\mathcal{J}_2(\varphi^{(\infty)})$	r
F	20	256	10^1	10^0	No closure	6.736×10^{-2}	8.876×10^{-3}	2.882×10^{-4}
G	20	512	10^{-1}	10^{-2}	No closure	6.736×10^{-2}	6.286×10^{-3}	1.685×10^{-4}

regularization performed in the present problem is less aggressive than in the problem discussed in Sec. V A. As shown in Fig. 5(a), the normalized error functional converges to a local minimum in only a few iterations, and, as the regularization is reduced, a larger reduction of the error functional is obtained. However, as is evident from Fig. 5(b), this is achieved with optimal eddy viscosities much better behaved than the optimal eddy viscosities found by solving the optimization problem discussed in Sec. V A, even though a weaker regularization is now applied (cf. Table II; the obtained
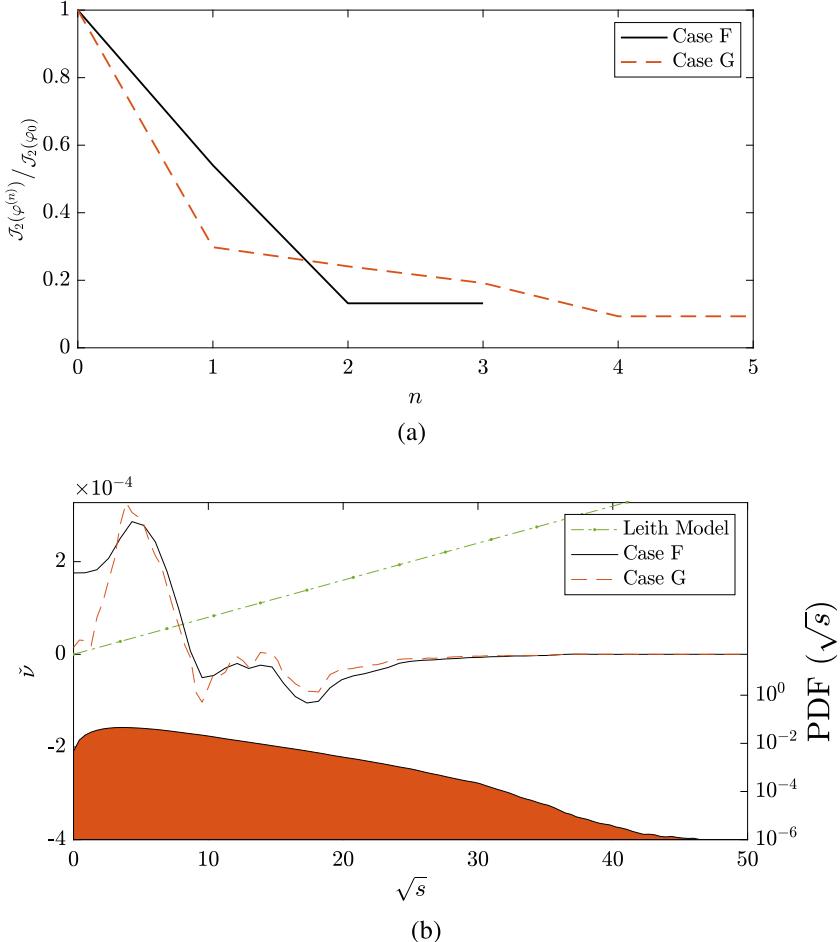

(a)

(b)

FIG. 5. (a) Dependence of the normalized functional $\mathcal{J}_2(\varphi^{(n)})/\mathcal{J}_2(\varphi_0)$ on the iteration n and (b) dependence of the corresponding optimal eddy viscosity $\check{\nu}$ on $\sqrt{s}$ for cases F and G; cf. Table II. Panel (b) also shows the PDF of $\sqrt{s}$ in case G.
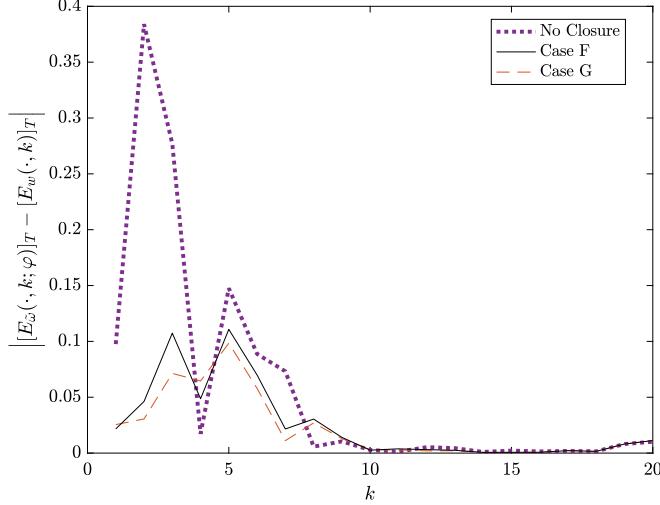
FIG. 6. The difference between time-averaged vorticity spectra (11) in the filtered DNS and in the LES with no closure and with the optimal eddy viscosity $\breve{\nu}$ obtained in cases F and G (cf. Table II) as function of the wave number k.

optimal eddy viscosity exhibits more small-scale variability in case G than in case F, but the difference is not significant).

The difference between the time-averaged vorticity spectra (11) is the LES with no closure, LES with the optimal closure $\breve{\nu}$ (cases F and G), and the filtered DNS is shown in Fig. 6 as a function of the wave number k [this quantity is related to the integrand expression in the error functional (12)]. We see that when the optimal eddy viscosity $\breve{\nu}$ is used in the LES, this error is reduced, especially at low wave numbers k. On the other hand, the evolution of the quantity $\log_{10}|1 - \mathcal{C}(t)|$ [cf. (28)], shown for the same cases in Fig. 7, demonstrates that, in contrast to Fig. 3, in the present problem the LES flows equipped with the optimal eddy viscosity do not achieve a better pointwise-in-space accuracy with respect to the DNS than the LES flow with no closure model.

Finally, we show the vorticity field $\widetilde{\omega}(T, \mathbf{x})$, the corresponding state variable $s(T, \mathbf{x})$ [cf. (7)], the spatial distribution $\breve{\nu}(s(T, \mathbf{x}))$ of the optimal eddy viscosity obtained in case G, and for comparison,
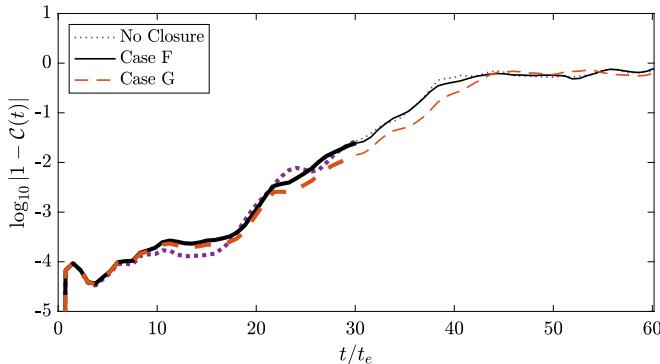

FIG. 7. Adjusted normalized correlations (28) for the LES with no closure and the optimal eddy viscosity in cases F and G. Thick and thin lines correspond to, respectively, time in the "training window" ($t \in [0, T]$) and beyond this window ($t \in (T, 2T]$).
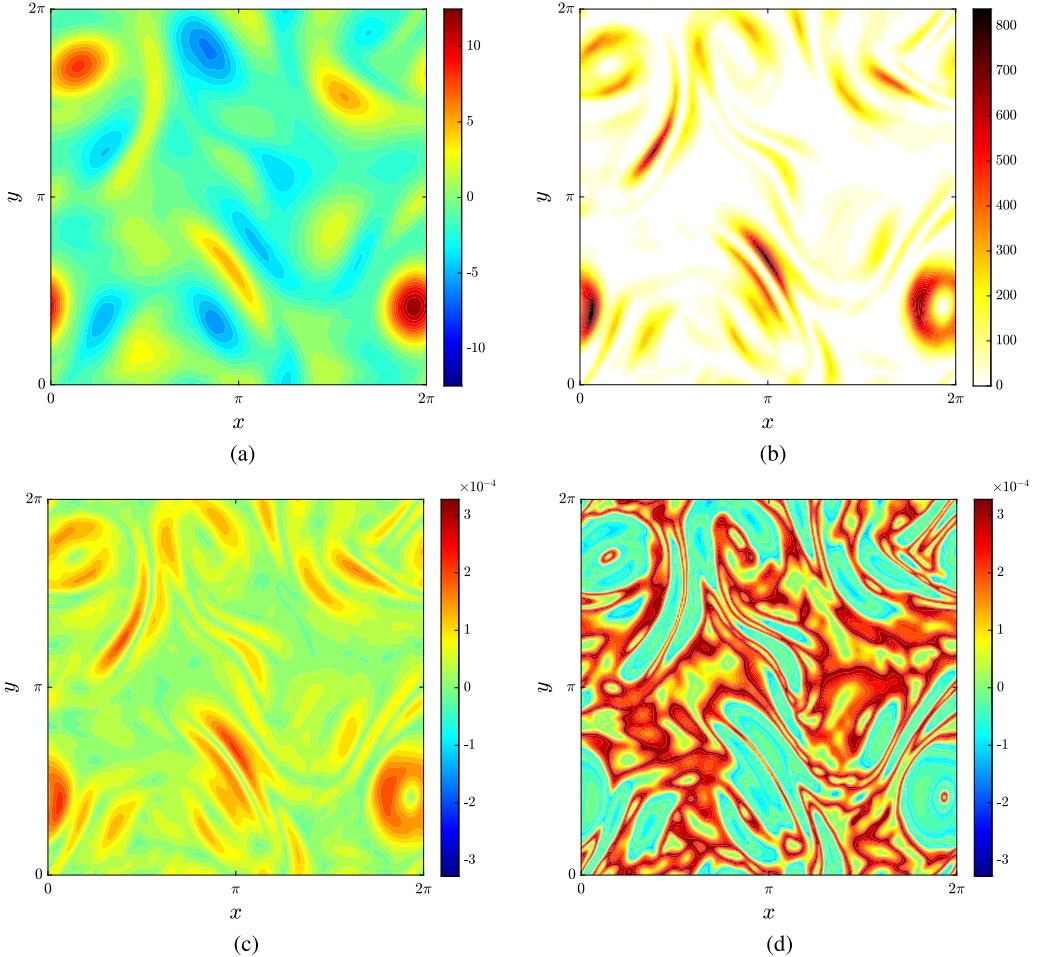
FIG. 8. For case G we show (a) the vorticity field $\widetilde{\omega}(T, \mathbf{x})$, $\mathbf{x} \in \Omega$, (b) the corresponding state variable $s(T, \mathbf{x})$ [cf. (7)], and the spatial distribution of (c) the eddy viscosity $\nu_L(s(T, \mathbf{x}))$ in the Leith model [cf. (6)] with $\delta = 0.02$, and (d) the optimal eddy viscosity $\breve{\nu}(s(T, \mathbf{x}))$ [cf. Fig. 5(b)], all shown at the end of the training window for $t = T$. For better comparison the same color scale is used in panels (c) and (d).

the spatial distribution of the eddy viscosity $\nu_L(s(T, \mathbf{x}))$ in the Leith model; cf. (6), in Figs. 8(a), 8(b), 8(d), and 8(c), respectively. We remark that the spatial distribution of the optimal eddy viscosity in Fig. 8(d) is now significantly smoother than the distribution of the optimal eddy viscosity obtained in the first formulation by solving optimization problem (14) with $j = 1$; cf. Fig. 4(d). An animated version of Fig. 8 illustrating the evolution of the fields for $t \in [0, 2T]$ is shown in the Supplemental Material as Movie 3 [38].

VI. DISCUSSION AND CONCLUSIONS

In this study we have considered the question of fundamental limitations on the performance of eddy-viscosity closure models for turbulent flows. We focused on the Leith model for 2D LES for which we sought optimal eddy viscosities that subject to minimum assumptions would result in the least mean-square error between the corresponding LES and the filtered DNS. Such eddy viscosities were found as minimizers of a PDE-constrained optimization problem with a nonstandard structure

which was solved using a suitably adapted adjoint-based gradient approach [17]. A key element of this approach was a regularization strategy involving the length-scale parameters ℓ_1 and ℓ_2 in the Sobolev gradients; cf. (26). The approach proposed is admittedly rather technically involved, which may limit its practical applicability to construct new forms of the eddy viscosity, but its value is in making it possible to systematically characterize the best possible performance of different types of closure models.

Our main finding in Sec. V A is that with a fixed cutoff wave number k_c the LES with an optimal eddy viscosity $\breve{\nu}$ matches the DNS increasingly well as the regularization in the solution of the optimization problem is reduced; cf. Fig. 2(a). This is quantified by a reduction of the rate of exponential decay of the correlation between the corresponding LES and the DNS; cf. Fig. 3(b) and Table I. This optimal performance of the closure model is achieved with eddy viscosities $\breve{\nu}(s)$ rapidly oscillating with a frequency increasing as the regularization parameters are reduced. From this we conclude that in the limit of vanishing regularization parameters and an infinite numerical resolution the optimal eddy viscosity would be undefined as it would exhibit oscillations with an unbounded frequency. Thus, from the mathematical point of view, the problem of finding an optimal eddy viscosity in the absence of regularization is ill-posed. In practical terms, this means that the "best" eddy viscosity for the Leith model does not exist.

The optimal performance of the LES is realized by a rapid variation of the eddy viscosity $\breve{\nu}(s)$ which oscillates between positive and negative values as s changes [cf. Fig. 2(b)], resulting in the dissipation and injection of the enstrophy occurring in the physical domain in narrow alternating bands; cf. Fig. 4(d). We note that a somewhat similar behavior was also observed in [14] where the authors used machine learning methods to determine pointwise estimates of eddy viscosity which exhibited oscillations between positive and negative values. This behavior can be understood in physical terms based on relations (4)–(5), which can be interpreted as defining the eddy viscosity in terms of the space- and time-dependent DNS field, but the problem is severely overdetermined. Thus, some form of relaxation is needed to determine ν, and the proposed optimization approach with its inherent regularization strategy is one possibility.

In addition, the optimal eddy viscosities found here have the property that $\breve{\nu}(0) > 0$, in contrast to what is typically assumed in the Leith model where $\nu(0) = 0$ [26]. In contrast to the behavior observed in Fig. 4(d), standard eddy viscosity closure models are usually assumed to be strictly dissipative [39], which is reflected in the fact that the eddy viscosity is nonnegative as in Fig. 4(c). We add that we have also considered finding optimal eddy viscosities by matching against the unfiltered DNS field, i.e., using $w(t, \mathbf{x})$ in the error functional (10) instead of $\widetilde{w}(t, \mathbf{x})$; however, this approach produced results very similar to the ones reported above. As is evident from Fig. 3(b), the performance of the LES with optimal eddy viscosities compares favorably to the LES with an optimal closure model proposed by Langford and Moser [37] based on a stochastic estimator, which has a less restrictive structure than the Leith model.

The optimal eddy viscosities constructed in Sec. V A to maximize the pointwise match against the filtered DNS are unlikely to be useful in practice due to their highly irregular behavior, which is difficult to resolve using finite numerical precision. On the other hand, the second formulation studied in Sec. V B where optimal eddy viscosities were determined by matching predictions of the LES against the time-averaged vorticity spectrum of the DNS for small wave numbers leads to a much better behaved optimization problem and produced results easier to interpret physically. In particular, the general form of the optimal eddy viscosity obtained in this case was found to have little dependence on regularization; cf. Fig. 5(b).

The main question left open by the results reported here is whether the optimal eddy viscosity for the Smagorinsky model in 3D turbulent flows would exhibit similar properties. It can be studied by solving an optimization problem analogous to (14), a task we will undertake in the near future. In addition, it is also interesting to analyze the optimal performance of other closure models using the framework developed here.

APPENDIX: GRADIENT OF THE ERROR FUNCTIONAL $\mathcal{J}_2$

Here we discuss computation of the gradients $\nabla_\varphi^{L^2} \mathcal{J}_2$ and $\nabla_\varphi^{H^2} \mathcal{J}_2$ of the error functional (12). The difference with respect to the formulation used in Sec. III is that functional (12) is defined in the Fourier space, and we adopt with suitable modifications the approach developed in [40]. Proceeding as in Sec. III, we first compute the Gâteaux differential of the error functional (12) with respect to φ,

$$\mathcal{J}_2'(\varphi;\varphi') = \frac{1}{2T} \int_{t=0}^{T} \int_{k=0}^{k_c} \left([E_{\widetilde{\omega}}(\cdot,k;\varphi)]_T - [E_w(\cdot,k)]_T \right) \left(\int_{\mathscr{C}(k)} \widehat{\widetilde{\omega}}\,\overline{\widehat{\widetilde{\omega}}}' + \overline{\widehat{\widetilde{\omega}}}\,\widehat{\widetilde{\omega}}'\, dS(\boldsymbol{k}) \right) dk\, dt, \quad (A1)$$

where $\overline{}$ denotes the complex conjugate and $\widehat{\widetilde{\omega}}'$ is the Fourier transform of the solution $\widetilde{\omega}'$ to (17). We note that the gradients $\nabla_\varphi^{L^2} \mathcal{J}_2$ and $\nabla_\varphi^{H^2} \mathcal{J}_2$ satisfy Riesz identities analogous to (19). Next we introduce new adjoint fields $\widetilde{\omega}^*$ and $\widetilde{\psi}^*$ assumed to satisfy the same adjoint system (22), but with a different source term W whose form is to be determined. Utilizing Parseval's identity and the fact that all fields are real-valued in physical space, we rewrite the duality relation (21) as

$$\left(\begin{bmatrix} \widetilde{\omega}' \\ \widetilde{\psi}' \end{bmatrix}, \mathcal{K}^* \begin{bmatrix} \widetilde{\omega}^* \\ \widetilde{\psi}^* \end{bmatrix} \right) = \frac{1}{2} \left(\widehat{\begin{bmatrix} \widetilde{\omega}' \\ \widetilde{\psi}' \end{bmatrix}}, \widehat{\mathcal{K}^* \begin{bmatrix} \widetilde{\omega}^* \\ \widetilde{\psi}^* \end{bmatrix}} \right) + \frac{1}{2} \overline{\left(\widehat{\begin{bmatrix} \widetilde{\omega}' \\ \widetilde{\psi}' \end{bmatrix}}, \widehat{\mathcal{K}^* \begin{bmatrix} \widetilde{\omega}^* \\ \widetilde{\psi}^* \end{bmatrix}} \right)},$$

$$= \frac{1}{2T} \int_{t=0}^{T} \int_{k=0}^{k_c} \int_{\mathscr{C}(k)} \widehat{\begin{bmatrix} \widetilde{\omega}' \\ \widetilde{\psi}' \end{bmatrix}} \cdot \overline{\widehat{\mathcal{K}^* \begin{bmatrix} \widetilde{\omega}^* \\ \widetilde{\psi}^* \end{bmatrix}}} + \overline{\widehat{\begin{bmatrix} \widetilde{\omega}' \\ \widetilde{\psi}' \end{bmatrix}}} \cdot \widehat{\mathcal{K}^* \begin{bmatrix} \widetilde{\omega}^* \\ \widetilde{\psi}^* \end{bmatrix}}\, dS(\boldsymbol{k})\, dk\, dt. \quad (A2)$$

Combining (17), (21), (22), (A1), and (A2) results in

$$\left(\begin{bmatrix} \widetilde{\omega}' \\ \widetilde{\psi}' \end{bmatrix}, \mathcal{K}^* \begin{bmatrix} \widetilde{\omega}^* \\ \widetilde{\psi}^* \end{bmatrix} \right) = \overbrace{\frac{1}{2T} \int_{t=0}^{T} \int_{k=0}^{k_c} \left([E_{\widetilde{\omega}}(\cdot,k;\varphi)]_T - [E_w(\cdot,k)]_T \right) \left(\int_{\mathscr{C}(k)} \widehat{\widetilde{\omega}}\,\overline{\widehat{\widetilde{\omega}}}' + \overline{\widehat{\widetilde{\omega}}}\,\widehat{\widetilde{\omega}}'\, dS(\boldsymbol{k}) \right) dk\, dt}^{\mathcal{J}_2'(\varphi;\varphi')},$$

from which we deduce the form of the source term in the adjoint system as

$$\widehat{W}(t,\boldsymbol{k}) = ([E_{\widetilde{\omega}}(\cdot,k;\varphi)]_T - [E_w(\cdot,k)]_T)\,\widehat{\widetilde{\omega}}(t,\boldsymbol{k}). \quad (A3)$$

Once the adjoint system (22) with the source term (A3) is solved, the L^2 gradient $\nabla_\varphi^{L^2} \mathcal{J}_2$ can be computed using expression (25). The Sobolev gradient $\nabla_\varphi^{H^2} \mathcal{J}_2$ is then obtained as discussed in Sec. III by solving system (26). In summary, the difference in the computation of the gradients of the error functionals $\mathcal{J}_1$ and $\mathcal{J}_2$ is confined to the form of the source term W in the adjoint system (22).

---

[1] M. Lesieur, *Turbulence in Fluids*, 2nd ed. (Kluwer, Dordrecht, 1993).

[2] S. B. Pope, *Turbulent Flows* (Cambridge University Press, Cambridge, 2000).

[3] P. A. Davidson, *Turbulence: An Introduction for Scientists and Engineers*, 2nd ed. (Oxford University Press, Oxford, 2015).

[4] J. N. Kutz, Deep learning in fluid dynamics, J. Fluid Mech. **814**, 1 (2017).

[5] M. Gamahara and Y. Hattori, Searching for turbulence models by artificial neural network, Phys. Rev. Fluids **2**, 054604 (2017).

[6] J. Jimenez, Machine-aided turbulence theory, J. Fluid Mech. **854**, R1 (2018).

[7] K. Duraisamy, G. Iaccarino, and H. Xiao, Turbulence modeling in the age of data, Annu. Rev. Fluid Mech. **51**, 357 (2019).

[8] K. Duraisamy, Perspectives on machine learning-augmented Reynolds-averaged and large eddy simulation models of turbulence, Phys. Rev. Fluids **6**, 050504 (2021).

[9] S. Pawar and O. San, Data assimilation empowered neural network parametrizations for subgrid processes in geophysical flows, Phys. Rev. Fluids **6**, 050501 (2021).

[10] J. Smagorinsky, General circulation experiments with the primitive equations: I. The basic experiment, Mon. Weather Rev. **91**, 99 (1963).

[11] C. E. Leith, Diffusion approximation for two-dimensional turbulence, Phys. Fluids **11**, 671 (1968).

[12] C. E. Leith, Atmospheric predictability and two-dimensional turbulence, J. Atmos. Sci. **28**, 145 (1971).

[13] C. E. Leith, Stochastic models of chaotic systems, Physica D **98**, 481 (1996).

[14] R. Maulik, O. San, and J. D. Jacob, Spatiotemporally dynamic implicit large eddy simulation using machine learning classifiers, Physica D **406**, 132409 (2020).

[15] V. Bukshtynov, O. Volkov, and B. Protas, On optimal reconstruction of constitutive relations, Physica D **240**, 1228 (2011).

[16] V. Bukshtynov and B. Protas, Optimal reconstruction of material properties in complex multiphysics phenomena, J. Comput. Phys. **242**, 889 (2013).

[17] P. Matharu and B. Protas, Optimal closures in a simple model for turbulent flows, SIAM J. Sci. Comput. **42**, B250 (2020).

[18] R. H. Kraichnan, Inertial ranges in two-dimensional turbulence, Phys. Fluids **10**, 1417 (1967).

[19] G. K. Batchelor, Computation of the energy spectrum in homogeneous two-dimensional turbulence, Phys. Fluids **12**, II-233 (1969).

[20] G. Boffetta, Energy and enstrophy fluxes in the double cascade of two-dimensional turbulence, J. Fluid Mech. **589**, 253 (2007).

[21] G. Boffetta and S. Musacchio, Evidence for the double cascade scenario in two-dimensional turbulence, Phys. Rev. E **82**, 016307 (2010).

[22] A. Bracco and J. C. McWilliams, Reynolds-number dependency in homogeneous, stationary two-dimensional turbulence, J. Fluid Mech. **646**, 517 (2010).

[23] A. Vallgren and E. Lindborg, The enstrophy cascade in forced two-dimensional turbulence, J. Fluid Mech. **671**, 168 (2011).

[24] G. Boffetta and R. E. Ecke, Two-dimensional turbulence, Annu. Rev. Fluid Mech. **44**, 427 (2012).

[25] J. P. Graham and T. Ringler, A framework for the evaluation of turbulence closures used in mesoscale ocean large-eddy simulations, Ocean Model. **65**, 25 (2013).

[26] R. Maulik and O. San, A dynamic framework for functional parameterisations of the eddy viscosity coefficient in two-dimensional turbulence, Int. J. Comut. Fluid Dyn. **31**, 69 (2017).

[27] R. A. Adams and J. F. Fournier, *Sobolev Spaces*, 2nd ed. (Elsevier/Academic Press, Amsterdam, 2003).

[28] B. Protas, T. R. Bewley, and G. Hagen, A computational framework for the regularization of adjoint analysis in multiscale PDE systems, J. Comput. Phys. **195**, 49 (2004).

[29] M. D. Gunzburger, *Perspectives in Flow Control and Optimization* (SIAM, Philadelphia, 2003).

[30] M. S. Berger, *Nonlinearity and Functional Analysis* (Academic Press, San Diego, 1977).

[31] J. Nocedal and S. Wright, *Numerical Optimization*, 2nd ed., Springer Series in Operations Research and Financial Engineering (Springer, New York, 2006).

[32] H. Le and P. Moin, An improvement of fractional step methods for the incompressible Navier-Stokes equations, J. Comput. Phys. **92**, 369 (1991).

[33] L. N. Trefethen, *Approximation Theory and Approximation Practice* (SIAM, Philadelphia, 2013).

[34] T. A. Driscoll, N. Hale, and L. N. Trefethen, *Chebfun Guide* (Pafnuty Publications, Oxford, 2014).

[35] T. Y. Hou, Blow-up or no blow-up? A unified computational and analytic approach to 3D incompressible Euler and Navier-Stokes equations, Acta Numer. **18**, 277 (2009).

[36] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes 3rd Edition: The Art of Scientific Computations* (Cambridge University Press, Cambridge, 2007).

[37] J. A. Langford and R. D. Moser, Optimal LES formulations for isotropic turbulence, J. Fluid Mech. **398**, 321 (1999).

[38] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevFluids.7.044605 for movies complementing Figs. 3(b), 4, and 8.

[39] W. Rodi, G. Constantinescu, and T. Stoesser, *Large-Eddy Simulation in Hydraulics* (CRC Press, Boca Raton, FL, 2013).

[40] M. M. Farazmand, N. K.-R. Kevlahan, and B. Protas, Controlling the dual cascade of two-dimensional turbulence, J. Fluid Mech. **668**, 202 (2011).