Online learning of eddy-viscosity and backscattering closures for geophysical turbulence using ensemble Kalman inversion

Yifei Guan [*]
Department of Mechanical Engineering, Union College, Schenectady, New York, USA and Department of Geophysical Sciences, The University of Chicago, Chicago, Illinois, USA

Pedram Hassanzadeh [†]
Department of Geophysical Sciences and Committee on Computational and Applied Mathematics, The University of Chicago, Chicago, Illinois, USA

Tapio Schneider and Oliver Dunbar
Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, California, USA

Daniel Zhengyu Huang
Beijing International Center for Mathematical Research, Peking University, Beijing, China

Jinlong Wu
Department of Mechanical Engineering, University of Wisconsin–Madison, Madison, Wisconsin, USA

Ignacio Lopez-Gomez
Google Research, Mountain View, California, USA

(Received 9 May 2025; accepted 3 May 2026; published 27 May 2026)

This article is part of a Physical Review *Collection on the Physics of Changing Climate.*

Different approaches to using data-driven methods for subgrid-scale closure modeling of geophysical turbulence have emerged recently. Most of these approaches are data hungry and lack interpretability and out-of-distribution generalizability. Here, we use a hybrid approach that combines turbulence theory, physics-based modeling, and data-driven methods to overcome these challenges. Specifically, we address the parametric uncertainty of well-known physics-based large-eddy simulation (LES) closures: the Smagorinsky (Smag) and Leith eddy-viscosity models (one free parameter) and the Jansen-Held (JH) backscattering model (two free parameters). For various cases of two-dimensional turbulence, optimal parameters are first learned online from data via ensemble Kalman inversion (EKI), such that for each case, the LES energy spectrum matches that of direct numerical simulation (DNS). We quantify the uncertainties on these parameters using a modern machine-learning-accelerated Bayesian workflow, "*Calibrate*, *Emulate*, *Sample*." Only a small training dataset is needed (to calculate the DNS spectra); i.e., the approach is data-efficient. We find the optimized parameter(s) and their associated uncertainty for each closure to be constant across broad flow regimes that differ in dominant length scales, eddy/jet structures, and dynamics, suggesting that these closures are generalizable. Next, we show that the online learned constants agree with the predictions of a recent semianalytical derivation, providing further interpretability. In both *a priori* and *a posteriori* tests that include examining the extreme events, LES with optimized closures, especially with JH, outperforms the baselines (LES with standard Smag, dynamic Smag, or Leith). This work shows the promise of combining advances in theory, physics-based modeling (e.g., JH), and data-driven modeling (e.g., online learning with EKI) to develop data-efficient frameworks for accurate, interpretable, and generalizable closures for geophysical turbulence, with ultimate applications in weather and climate prediction.

DOI: 10.1103/mnbm-3g56

I. INTRODUCTION

Climate models run at coarse resolutions (10–100 km grid spacing) and therefore cannot directly resolve many

[*]Contact author: guany@union.edu
[†]Contact author: pedramh@uchicago.edu

critical processes such as turbulence and convection. To model the small-scale processes, climate models rely on large-eddy simulations (LES). The stability and accuracy of LES depend on the quality of the subgrid-scale (SGS) model. These closures, sometimes also referred to as parametrizations, represent the effects of unresolved (subgrid) processes on the resolved (large-scale) flow in LES. For example, based on the Boussinesq hypothesis, Smagorinsky proposed an eddy-viscosity closure model for numerical weather prediction that treats the SGS effects as pure diffusion [1]. A similar eddy-viscosity type model was proposed later by Leith [2,3] to model the two dimensional (2D) or quasi-2D turbulence similar to the large-scale atmospheric and oceanic flows. Although the eddy-viscosity models succeed in ensuring stability, they cannot capture the backscattering effect, which is the energy transferred from small (subgrid) to large scales. The backscattering effect is important in oceanic and atmospheric circulations [4–6]. Therefore, SGS models with backscattering have been proposed to increase the accuracy while maintaining stability [7–10]. However, such SGS models often involve free parameters that are chosen empirically or on a case-by-case basis, also known as "parametric uncertainty."

Recent advances in data-driven methods offer new avenues to develop better SGS closures (e.g., Refs. [11–17]). Most studies have focused on *offline* (supervised) learning, in which a closure is developed, for instance, by training neural networks or using equation-discovery algorithms to match the true and parametrized SGS fluxes (e.g., Refs. [18–23]). The trained/discovered closure, once validated *a priori*, is coupled to a low-resolution solver (e.g., a climate model) for *a posteriori* tests. While examples of stable and accurate SGS modeling using *offline* learning with neural networks have emerged recently (e.g., Refs. [24–27]), this approach has major drawbacks: (1) training neural networks requires many samples of the "true" SGS fluxes, (2) these closures are largely uninterpretable, and (3) they do not generalize out-of-distribution (extrapolate) to other flows (e.g., higher Reynolds numbers, Re) (see, e.g., Refs. [28–30]). Item (1) is a major practical challenge as it requires long, expensive, high-resolution simulations (e.g., direct numerical simulations, DNS). Also, estimating the true SGS fluxes from such data is difficult [22,31].

An alternative is *online* (also referred to as "end-to-end," "*a posteriori,*" or "on-the-fly") learning: The closure is developed while it is coupled to the low-resolution numerical solver, with the objective of matching certain statistics of the "true" flow (e.g., from high-resolution simulations or even observations) (e.g., Refs. [32,33]). If the existing closures have "structural uncertainties," then neural networks can be used to learn either the entire closure from data (e.g., Ref. [34]) or corrections to physics-based closures (e.g., Ref. [35]). However, minimizing the loss function of this approach can be practically challenging (see below); that said, promising results with neural networks trained using ensemble Kalman inversion (EKI) [36], multiagent reinforcement learning [37,38], ensemble-based synchronization [39,40], and differentiable modeling [41–44] have started to appear. Although this approach addresses item (1), interpretability and generalization [(2) and (3)] remain major challenges.

One avenue for addressing (1)–(3) is to develop closures with the appropriate structure using physical arguments (and when needed, equation-discovery techniques [23]), and use *online* learning to estimate the unknown parameters (i.e., addressing parametric uncertainty); see Ref. [45] for further discussions and Refs. [12,46–49] for recent examples of success with this approach. A major question of this approach is whether calibrating/tuning a closure for one limited set of metrics generalizes to performance for other metrics. Furthermore, generalization (3) requires the estimated parameters to be universal constants (or at least nearly invariant within the flow regimes of interest). These questions, even for canonical test cases, remain to be fully investigated. *They are the subject of the current study*.

To achieve quantitative parametric uncertainty in the context of modern learning algorithms, we embed an EKI optimizer within an approximate parametric uncertainty pipeline known as "*Calibrate*, *Emulate*, *Sample*" (CES) [50]. This three-stage process first applies an ensemble Kalman method (here, we use EKI). The EKI is not just used for optimizing parameters; instead, all evaluations are collected and used as training pairs for an emulator (here, we use a Gaussian Process) (e.g., Refs. [33,51,52]). Finally, this emulator is fed into a sampling method [here, we use the Markov Chain Monte Carlo (MCMC) method] for parametric uncertainty quantification. The motivation of such an algorithm is to overcome the well-known underestimation of uncertainty in ensemble optimizer approaches ("ensemble collapse"), to reduce the high computational cost of sampling and to ensure good accuracy of the emulator during sampling. It should be mentioned that, though specific algorithmic choices must be made, generalization to other ensemble optimizers (e.g., Kalman variants, hybrid ensemble-variational methods, etc.), statistical emulators, or samplers, can be made. CES has been successfully applied in the context of Earth system modeling to develop a variety of closures online [11,53–56].

In this paper, we aim to answer the interpretability and generalization questions of SGS models for seven setups of 2D homogeneous isotropic turbulence, plus one that includes a planetary vorticity gradient (β-plane effect), which are canonical models for many geophysical flows (e.g., Refs. [57–60]). We estimate the parameters of two classical eddy-viscosity closures (Smagorinsky and Leith) and a new backscattering closure [8] *online* using EKI and by matching the energy spectrum of LES with that of a short DNS (truth). We use the EKI iterations within CES to further learn the parametric uncertainty. The LES with EKI-optimized parameters, which are found to agree with a recent semianalytical model for these parameters [61], are then examined in terms of enstrophy spectra, interscale transfers, and probability density functions of vorticity, especially at the tails (extreme events).

II. DATA AND METHODS

A. DNS and LES of 2D turbulence

We use seven setups of 2D homogeneous isotropic turbulence, plus one that includes a planetary vorticity gradient (β-plane effect). The dimensionless governing equations in the vorticity (ω) and streamfunction (ψ) formulation in a

TABLE I. Top: Physical and numerical parameters of the total 8 cases. Bottom: EKI-optimized and semianalytically-derived parameters (uncertainties are shown in Table II). The value of A is diagnosed from the inertial region of the DNS turbulent kinetic energy (TKE) spectrum for each case. The semianalytical values $C_{\mathrm{S}}^{\mathrm{Analytical}}$, $C_{\mathrm{L}}^{\mathrm{Analytical}}$, and $C_{\mathrm{JHL}}^{\mathrm{Analytical}}$ are given by Eqs. (6), (7) and (8), and $C_{\mathrm{B}}^{\mathrm{Analytical}}$ is chosen to be the same as $C_{\mathrm{B}}^{\mathrm{EKI}}$ for JHL for direct comparison.

Case	1.1	1.2	1.3	1.4	2	3.1	3.2	3.3
Re	20 000	20 000	100 000	300 000	20 000	20 000	100 000	300 000
k_f	4	4	4	4	4	25	25	25
β	0	0	0	0	20	0	0	0
N_{DNS}	1024	1024	4096	4096	1024	1024	4096	4096
N_{LES}	32	64	256	256	64	256	256	256
A	1.87	1.87	1.85	1.81	2.48	1.88	1.77	1.79
$C_{\mathrm{S}}^{\mathrm{EKI}}$	0.12	0.12	0.11	0.12	0.10	0.12	0.12	0.10
$C_{\mathrm{S}}^{\mathrm{Analytical}}$	0.13	0.12	0.11	0.12	0.10	0.11	0.12	0.12
$C_{\mathrm{L}}^{\mathrm{EKI}}$	0.23	0.25	0.26	0.24	0.21	0.24	0.23	0.21
$C_{\mathrm{L}}^{\mathrm{Analytical}}$	0.23	0.23	0.23	0.24	0.20	0.23	0.24	0.24
$C_{\mathrm{JHS}}^{\mathrm{EKI}}$, $C_{\mathrm{B}}^{\mathrm{EKI}}$	0.23, 0.96	0.22, 0.95	0.21, 0.95	0.21, 0.95	0.20, 0.94	0.22, 0.94	0.23, 0.95	0.21, 0.96
$C_{\mathrm{JHL}}^{\mathrm{EKI}}$, $C_{\mathrm{B}}^{\mathrm{EKI}}$	0.34, 0.95	0.32, 0.95	0.33, 0.94	0.30, 0.94	0.31, 0.96	0.32, 0.95	0.32, 0.94	0.31, 0.93
$C_{\mathrm{JHL}}^{\mathrm{Analytical}}$, $C_{\mathrm{B}}^{\mathrm{Analytical}}$	0.35, 0.95	0.34, 0.95	0.33, 0.94	0.33, 0.94	0.31, 0.96	0.33, 0.95	0.34, 0.94	0.33, 0.93

doubly periodic square domain of length $L = 2\pi$ are

$$\frac{\partial \omega}{\partial t} + \mathcal{N}(\omega, \psi) = \frac{1}{Re}\nabla^2\omega - f - r\omega + \beta\frac{\partial \psi}{\partial x}, \quad \nabla^2\psi = -\omega. \tag{1}$$

Here, $\mathcal{N}$ is the Jacobian, $f(x, y) = k_f[\cos(k_f x) + \cos(k_f y)]$ is a deterministic forcing, and $r = 0.1$ is the linear drag coefficient (e.g., representing surface friction); β is the gradient of the Coriolis parameter. We study eight cases, in which Reynolds number (Re), forcing wavenumber (k_f), and β have been varied (Table I), creating a variety of flows that differ in dominant length scales and energy/enstrophy cascade regimes (see Fig. 1). For DNS, which is treated as the truth, Eq. (1) is solved at high spatial-temporal resolutions ($N_{\mathrm{DNS}}^2 \times \Delta t_{\mathrm{DNS}}$, where N_{DNS} is the DNS resolution in one direction and Δt_{DNS} is the DNS simulation time step size), using a Fourier pseudospectral solver (see Ref. [21]). The LES equations are derived by applying a low-pass filter (e.g., sharp cut-off), denoted by $\overline{(\cdot)}$, to Eq. (1):

$$\begin{aligned}\frac{\partial \overline{\omega}}{\partial t} + \mathcal{N}(\overline{\omega}, \overline{\psi}) =& \frac{1}{Re}\nabla^2\overline{\omega} - \overline{f} - r\overline{\omega} + \beta\frac{\partial \overline{\psi}}{\partial x} \\ &- \underbrace{\left[\overline{\mathcal{N}(\omega, \psi)} - \mathcal{N}(\overline{\omega}, \overline{\psi})\right]}_{\Pi^{\mathrm{SGS}} = \nabla\times(\nabla\cdot\tau^{\mathrm{SGS}})}, \\ \nabla^2\overline{\psi} =& -\overline{\omega}.\end{aligned} \tag{2}$$

Solving these equations needs much coarser resolutions ($N_{\mathrm{LES}}^2 \times 10\Delta t_{\mathrm{DNS}}$, where N_{LES} is the LES resolution in one direction). In this work, $N_{\mathrm{DNS}}/N_{\mathrm{LES}}$ is between 4 and 32 (Table I). The SGS term, Π^{SGS} or τ^{SGS}, requires a closure.

Physics-based closures with parametric uncertainty

Eddy-viscosity closures are "functional" models that assume the net effect of the SGS term is dissipation of the resolved scales: $\tau^{\mathrm{SGS}} = -2\nu_e\bar{\mathcal{S}}$, where $\bar{\mathcal{S}}$ is the rate of strain of the resolved flow and $\nu_e(x, y, t)$ is the eddy viscosity [62]. For 2D turbulence, the Smagorinsky model [Smag: $\nu_e = (C_{\mathrm{S}}\Delta)^2\langle\bar{\mathcal{S}}^2\rangle^{1/2}$, $\Delta = L/N_{\mathrm{LES}}$] [1] and the Leith model [Leith: $\nu_e = (C_{\mathrm{L}}\Delta)^3\langle(\nabla\bar{\omega})^2\rangle^{1/2}$] [3] are often used in practice, with C_{S} and C_{L} being model parameters to be determined ($\langle\cdot\rangle$ means domain averaging). For 3D homogeneous isotropic turbulence, a constant $C_{\mathrm{S}} = 0.17$ was analytically derived from the energy spectrum scaling [63,64]. Empirically, it was found that 0.17 leads to the best *a priori* performance [65]. For 2D turbulence, no analytical or empirical estimate existed for C_{S} or C_{L}, and trial and error was often used to select a value [18] until very recently, when a semianalytical estimate for C_{S} and C_{L} was presented in Ref. [61]. For both closures, C can be also determined dynamically (DSmag, DLeith) from the Germano identity [6,66,67]. However, this procedure can lead to $\nu_e < 0$ (backscattering) and unstable LES. While capturing backscattering is desirable (see below), in practice, DSmag and DLeith are used with "positive clipping," which enforces $\nu_e \geqslant 0$ and makes the closure more diffusive (thus less accurate) in favor of stability.

To capture backscattering, which can be important in many flows [62], including oceanic and atmospheric circulations [4–6], backscattering closures have been developed that, for example, inject a fraction of the dissipated energy back to the resolved scales. Here, we use the recently developed closure of Jansen and Held [8] (JH, hereafter), which has been applied to oceanic and quasigeostrophic turbulence [9,26]. The JH closure is

$$\begin{aligned}\Pi^{\mathrm{SGS}} &= \nabla^2(\nu_e\nabla^2\bar{\omega}) + \nu_b\nabla^2\bar{\omega}, \\ \nu_b &= -C_{\mathrm{B}}\langle\bar{\psi}\nabla^2(\nu_e\nabla^2\bar{\omega})\rangle/\langle\bar{\psi}\nabla^2\bar{\omega}\rangle,\end{aligned} \tag{3}$$

where the first term is biharmonic eddy viscosity and the second term represents backscattering with antidiffusion. C_{B} determines how much of the globally dissipated energy is fed back into the resolved scales. ν_e can be determined in the same way as in Leith [8] or Smag [9,26]. Like C_{S} and C_{L}, C_{B} is an empirically determined parameter that is expected to be $\mathcal{O}(1)$ to balance the forward transferred and backscattered energy.

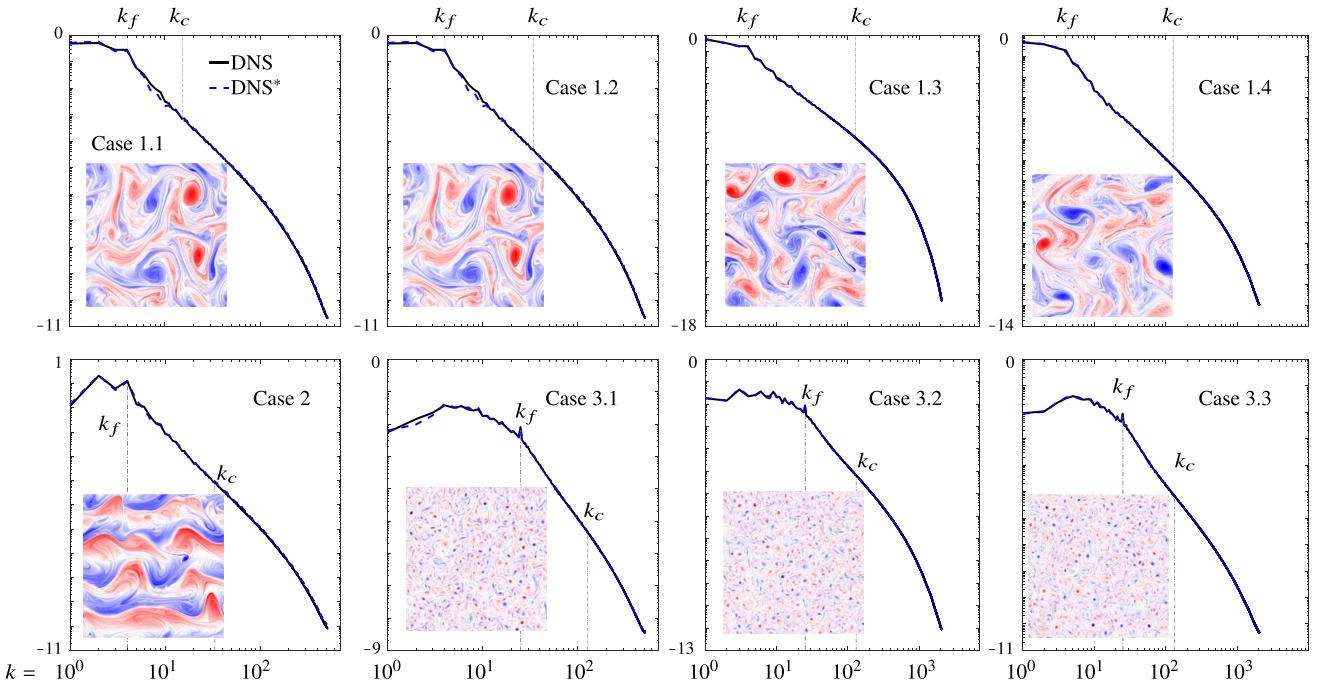

FIG. 1. The TKE spectra ($\log_{10}(\hat{E}(k))$) and examples of ω for each case. The forcing wavenumber k_f and cut-off wavenumber k_c are marked. The black (blue) lines show DNS spectra averaged over all 100 (first five) snapshots in the training set. With a sharp-spectral cut-off filter, the FDNS spectrum, which is used as the target in Eq. (4), overlaps with the DNS spectrum over $k = [0, k_c]$.

JH showed that in general, $C_B \geqslant 0.9$, and used 0.9 in their research [8].

Here, we focus on four closures: Smag, Leith, JH-Smag (JHS), and JH-Leith (JHL). We estimate, as described below, their one or two *constant* parameters C for each case. For JHS and JHL, we define ν_e in similar ways as in Smag/Leith [8,9,26] but choose the power of $C\Delta$ consistent with the dimension of the biharmonic: $\nu_e = (C_{\text{JHS}}\Delta)^4\langle\bar{\mathcal{S}}^2\rangle^{1/2}$ (JHS) and $\nu_e = (C_{\text{JHL}}\Delta)^6\langle(\nabla^2\bar{\omega})^2\rangle^{1/2}$ (JHL). Note that to be consistent with other implementations, unlike JH14 or JH15, we use domain averaging in the calculation of ν_e^{JHL}. Here, like C_S and C_L, parameters C_{JHS}, C_{JHL}, and C_B are dimensionless.

B. Calibrate, emulate, sample for *online* learning

The full CES framework applied in this work is depicted in Fig. 2. We estimate the optimal C parameter(s) (e.g., C_S, C_L, C_{JHS}, C_{JHL}, C_B) of a closure by matching the TKE spectra ($\hat{E}(k)$) of the filtered DNS (FDNS) and of the LES with that closure (we have also explored matching enstrophy spectra $\hat{Z}(k) = k^2\hat{E}(k)$ (see Sec. IV). More specifically, we minimize the following loss function with respect to the C value(s):

$$\mathcal{L} = \left\{\ln(\hat{E}^{\text{FDNS}}) - \ln(\hat{E}^{\text{LES}}),\ \Gamma_{\text{FDNS}}^{-1}(\ln(\hat{E}^{\text{FDNS}}) - \ln(\hat{E}^{\text{LES}}))\right\}, \tag{4}$$

where $\{,\}$ is the Euclidean inner product (dot product of two vectors whose elements are indexed by wavenumber k), Γ_{FDNS} is the covariance between wavelengths in $\ln(\hat{E}^{\text{FDNS}})$, and $k = (k_x^2 + k_y^2)^{1/2}$ is the wavenumber.

To obtain the FDNS data and to construct the training and testing datasets, we apply a low-pass sharp-spectral filter with cut-off wavenumber $k_c = L/(2\Delta) = N_{\text{LES}}/2$ to each DNS snapshot. Therefore, for each case, the FDNS spectrum matches that of the DNS up to k_c (Fig. 1). Here, we use the sharp-spectral filter for two reasons: (1) the FDNS spectrum is exactly the same as the DNS spectrum up to k_c, and therefore the EKI-optimized parameters can be compared with our new semianalytical derivation [61] (also see Table I), and (2) *offline* learning has been found unable to learn from FDNS data with the sharp-spectral filter [21,24], which is a key advantage of *online* vs *offline* learning in SGS modeling. It should be noted that the choice of filtering is not exclusive. *Online* learning with CES can work with training data obtained from a Gaussian filter or a Box filter. In such cases, however, the EKI-optimized and the semianalytically-derived parameters are expected to be different than the ones obtained in this work using the sharp-spectral filter. $\hat{E}(k)$ and Γ_{FDNS} in Eq. (4), the only information from DNS needed for training, are calculated by averaging the $\hat{E}(k)$ profiles over 100 decorrelated FDNS snapshots to obtain a reliable estimate (see the Supplemental Material for more details [68]). The DNS length for this number of FDNS snapshots is within the "small data" regime for these cases and not enough for *offline* training of an accurate and stable network-based closure [*offline* training typically requires $\mathcal{O}(10^4)$ DNS snapshots] [21,24]. The computational costs comparison between DNS, LES, and EKI is included in Table III of the Supplemental Material [68]. Note that because $\hat{E}(k)$ is rather invariant for flows in statistical equilibrium, we could have used an even shorter DNS from each case, as $\hat{E}(k)$ spectra averaged over just the first 5 DNS snapshots (i.e., a 20× shorter DNS) are virtually indistinguishable from those calculated from all 100 snapshots (see Fig. 1), and the optimized parameters are within 1% difference. If other statistics, such as intermittency, were used as the optimization target, they may not remain the same with a smaller

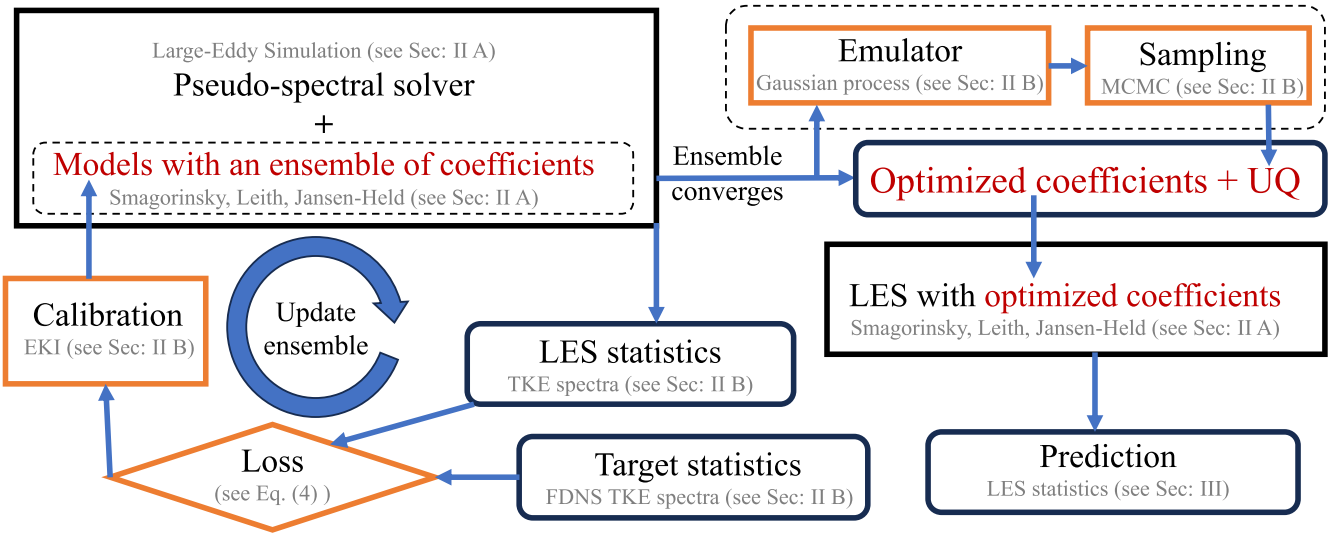

FIG. 2. The schematic of the CES framework with EKI to estimate the closure parameters with uncertainty. Briefly, first, for a given case and closure, an ensemble of C values is generated by drawing 10 times from a Gaussian distribution (see the Supplemental Material [68]). For each C value(s), a long-enough LES is performed and the mean TKE spectrum of this ensemble member is computed. Then the loss is calculated against the target (FDNS spectrum) for each ensemble member. EKI is used to update the C values, and this process continues until the ensembles converge, which usually happens in 10 or fewer epochs (a typical convergence trend is shown in Fig. 1 of the Supplemental Material [68]). For uncertainty quantification, all parameters and their corresponding statistics from all epochs are used to train a Gaussian process emulator. This fast emulator is used for uncertainty quantification with MCMC.

dataset. However, here we find that the energy spectra are sufficient to learn closures that capture many other statistical properties of the flow, including those of the rare, extreme events of vorticity. Whether these findings hold for more complex systems and other variables remains to be investigated (though the closure model only provides coarse-grained feedback on the large-scale resolved flow).

Minimizing loss functions such as Eq. (4) requires performing LES during the optimization process. While traditional data assimilation methods, such as four-dimensional-Variational algorithms (e.g., Ref. [69]), rely on gradient descent, which requires the explicit calculation of an adjoint model, derivative-free techniques such as EKI [70–72] can be used with any solver, which may not be differentiable. In this type of approach, optimizing the model parameters C is framed as maximizing the *posterior* probability, obtained by combining prior assumptions with data:

$$P(C \mid \ln(\hat{E}^{\mathrm{FDNS}}), \Gamma_{\mathrm{FDNS}}) = \frac{e^{-\mathcal{L}}}{\zeta(\ln(\hat{E}^{\mathrm{FDNS}}), \Gamma_{\mathrm{FDNS}})} P_{\mathrm{prior}}(C), \tag{5}$$

where $\mathcal{L}$ is the loss as in Eq. (4), $\zeta(\ln(\hat{E}^{\mathrm{FDNS}}), \Gamma_{\mathrm{FDNS}})$ is a normalizing constant, and $P_{\mathrm{prior}}(C)$ is the *prior* probability density of C (initial assumptions about the possible values of C). Here, as $\mathcal{L}$ is a weighted quadratic term, the distribution of $e^{-\mathcal{L}}$ is Gaussian.

To a broader discussion, EKI is one of the family of ensemble Kalman methods, which are in turn part of a family of ensemble data-assimilation methods that have been repurposed to solve inverse problems. In related investigations [73–75], a hybrid ensemble-variational method [76] was successfully applied for learning turbulence closure parameters. In practice, ensemble data-assimilation methods may differ in applicability or scalability, but across many problems may give similar performance, as observed, for example in Ref. [76]. While we have used EKI, we could have used any data-assimilation-based approach; note that these approaches are preferred over alternative classes of derivative-free methods such as search methods (e.g., grid search, random search [77]) or metaheuristics [78], which are well known to require high numbers of function evaluations and suffer from curse of dimensionality (e.g., Ref. [79]).

Unlike most other ensemble data-assimilation methods, ensemble Kalman approaches also enjoy a quickly growing mathematical literature (e.g., Ref. [80]) that has led to several improved algorithmic variants and features seen in modern software packages (e.g., Ref. [81]), including adaptive time stepping, regularization, sampling-error correction, and acceleration, to improve convergence and reduce computational resources, which could be important for scaling up to complex weather and climate models.

In addition to optimization/calibration, we are also interested in quantifying the uncertainty of the *posterior* distribution [Eq. (5)]. In Ref. [73], the authors used the empirical covariance of the particles to approximate the covariance of the *posterior* distribution of C. We did not follow this approach for two reasons: (1) EKI particle spread is often overconfident ("ensemble collapse") and therefore is not suitable for uncertainty quantification without careful inflation procedures (e.g., ensemble Kalman sampler [82] or affine invariant interacting Langevin dynamics [83]), and (2) ensemble samplers typically require orders of magnitude more LES runs to converge than ensemble optimizers.

Instead, we obtain an approximate uncertainty quantification by following the approach of CES [50], which performs sampling (with the MCMC method) using an emulator trained on the EKI trajectories (all LES results during the EKI train-

TABLE II. Uncertainty of EKI-optimized parameters for all eight cases is obtained using the MCMC method in the CES framework.

Case	1.1	1.2	1.3	1.4
$C_{\mathrm{S}}^{\mathrm{EKI}}$	0.012	0.010	0.0041	0.0082
$C_{\mathrm{L}}^{\mathrm{EKI}}$	0.032	0.028	0.028	0.025
$C_{\mathrm{JHS}}^{\mathrm{EKI}}$, $C_{\mathrm{B}}^{\mathrm{EKI}}$	0.018, 0.012	0.010, 0.010	0.0044, 0.013	0.0074, 0.013
$C_{\mathrm{JHL}}^{\mathrm{EKI}}$, $C_{\mathrm{B}}^{\mathrm{EKI}}$	0.019, 0.012	0.0095, 0.010	0.0043, 0.013	0.0064, 0.013
Case	2	3.1	3.2	3.3
$C_{\mathrm{S}}^{\mathrm{EKI}}$	0.012	0.008	0.011	0.015
$C_{\mathrm{L}}^{\mathrm{EKI}}$	0.015	0.026	0.024	0.035
$C_{\mathrm{JHS}}^{\mathrm{EKI}}$, $C_{\mathrm{B}}^{\mathrm{EKI}}$	0.0096, 0.012	0.0039, 0.012	0.0057, 0.012	0.0024, 0.012
$C_{\mathrm{JHL}}^{\mathrm{EKI}}$, $C_{\mathrm{B}}^{\mathrm{EKI}}$	0.014, 0.013	0.0051, 0.011	0.0036, 0.013	0.0099, 0.0098

ing process). The emulator both accelerates and smooths the sampling, and is trained on well-chosen points around the *posterior* mode (all the non-blow-up LES data obtained from the EKI process; see e.g., Fig. 1 in the Supplemental Material [68]). During this stage, no further LES runs are required.

In addition to TKE spectra ($\hat{E}(k)$), other statistics, such as mean shear stresses, can also be used as targets [73] in the loss function [Eq. (4)]. Here, we find that using TKE spectra as the target provides the best performance and interpretation as it is directly related to our newly proposed semianalytical derivation [61]. To quantify the uncertainty of the estimated C value(s), we employ the CES framework [50]. The framework is well established and now available as a software [84].

III. RESULTS

A. Online-learned parameters

Table I shows the EKI-optimized constant parameters of Smag, Leith, JH-Smag, and JH-Leith for all eight cases. Table II shows the corresponding uncertainty for each parameter. The reported values in Table II are the posterior uncertainties associated with the estimated parameters, obtained from the Markov-Chain Monte Carlo sampling procedure. These values correspond to the standard deviations (credible intervals) of the posterior distributions. The $C_{\mathrm{S}}^{\mathrm{EKI}}$ is around 0.12 and the same for all cases [within the $\mathcal{O}(0.01)$ uncertainty]. This value is smaller than the analytical and empirical value of 0.17 often used based on previous work with 3D homogeneous isotropic turbulence. As shown below, the EKI-optimized value leads to significant improvements in *a priori* and *a posteriori* metrics over Smag with $C_{\mathrm{S}} = 0.17$ (Smag-0.17). The $C_{\mathrm{L}}^{\mathrm{EKI}}$ is around 0.24, and again, the same for all cases [within the estimated uncertainty $\mathcal{O}(0.01)$]. $C_{\mathrm{JHS}}^{\mathrm{EKI}}$ and $C_{\mathrm{JHL}}^{\mathrm{EKI}}$ are also found to be similar across all cases and around 0.215 and 0.32, respectively. The backscattering parameter $C_{\mathrm{B}}^{\mathrm{EKI}}$ is the same across all cases and around 0.95, which is between unity and the previously used value of 0.9. In summary, across these eight cases, C_{S}, C_{L}, C_{JHS}, C_{JHL}, and C_{B} in these four closures are found to be nearly constant. Later in Sec. III B, we compare these parameters with those estimated from a semianalytical derivation [61], which sheds light on this near universality in the constant value.

For baselines in the *a priori* and *a posteriori tests*, we conduct LES with Π^{SGS} parametrized using Smag-0.17 and using DSmag and DLeith (with positive clipping) implemented following recent studies [24,85]. LES with these three "baseline" closures and with the four EKI-optimized closures are conducted for each case. Each LES run is 100× longer than the duration of the DNS used for the training dataset. As discussed below, for *a priori* tests, we compare the total energy and enstrophy interscale transfers among these closures and FDNS. For *a posteriori* tests, we compare the enstrophy spectra and vorticity probability density functions from LES with different closures and from FDNS.

The global energy and enstrophy transfers are $\langle P_E \rangle = \langle \Pi^{\mathrm{SGS}} \bar{\psi} \rangle$ and $\langle P_Z \rangle = \langle \Pi^{\mathrm{SGS}} \bar{\omega} \rangle$, respectively [24,86]. Here, $\langle P_E \rangle$ or $\langle P_Z \rangle$ is positive for forward transfer (energy/enstrophy moving to SGS from resolved scales) and negative for backscatter (moving to resolved scales from SGS). $\langle P_E \rangle$ and $\langle P_Z \rangle$ from FDNS and LES for representative cases (1.1, 1.4, 2, 3.3) are shown in Fig. 3 (values for all cases are reported in Supplemental Material [68]). It can be observed that for all cases, JHS and JHL have the best agreement with FDNS, and for half of the cases, the values of $\langle P_E \rangle$ and $\langle P_Z \rangle$ closely match those of FDNS. Exceptions are Cases 1.2, 1.3, 1.4, and 3.1, for which JHS and JHL underpredict $\langle P_E \rangle$ by a factor of 2–3; however, they still significantly outperform the optimized Smag and Leith and the baselines, which overpredict $\langle P_E \rangle$ by 1 and even 2 orders of magnitude (thus, these closures are too diffusive). The gain from optimization can be isolated by comparing Smag and Smag-0.17: The former, which uses $C_{\mathrm{S}} \approx 0.12$, consistently outperforms the baseline in both $\langle P_E \rangle$ and $\langle P_Z \rangle$ (often by a factor of 2–5, or even larger). Error bars are included in Fig. 3 for EKI-optimized models to show the credible intervals. The uncertainties in $\langle P_E \rangle$ and $\langle P_Z \rangle$ predictions result from the uncertainties in EKI-optimized parameters (Table II). Next, we examine the enstrophy spectra, $\hat{Z}(k)$, and the probability density function of vorticity, $\mathcal{P}(\omega)$. Figure 4 shows $\hat{Z}(k)$ for representative cases. Smag-0.17 and DSmag, consistent with $\langle P_Z \rangle$ in Fig. 3, have excessive enstrophy dissipation and underpredict, by about an order of magnitude, $\hat{Z}(k)$ compared to FDNS at the smallest scales. Among the baselines, DLeith performs the best in matching the FDNS, but it is still not comparable to EKI-optimized models. All the EKI-optimized closures work well in matching the FDNS $\hat{Z}(k)$, which is not surprising given that their target was TKE spectra [$\hat{E}(k)$]. The one exception is Case 3.3, where only optimized JHS and JHL closures (and to a lesser degree Leith) can match the FDNS $\hat{Z}(k)$ in the largest scales (small k).

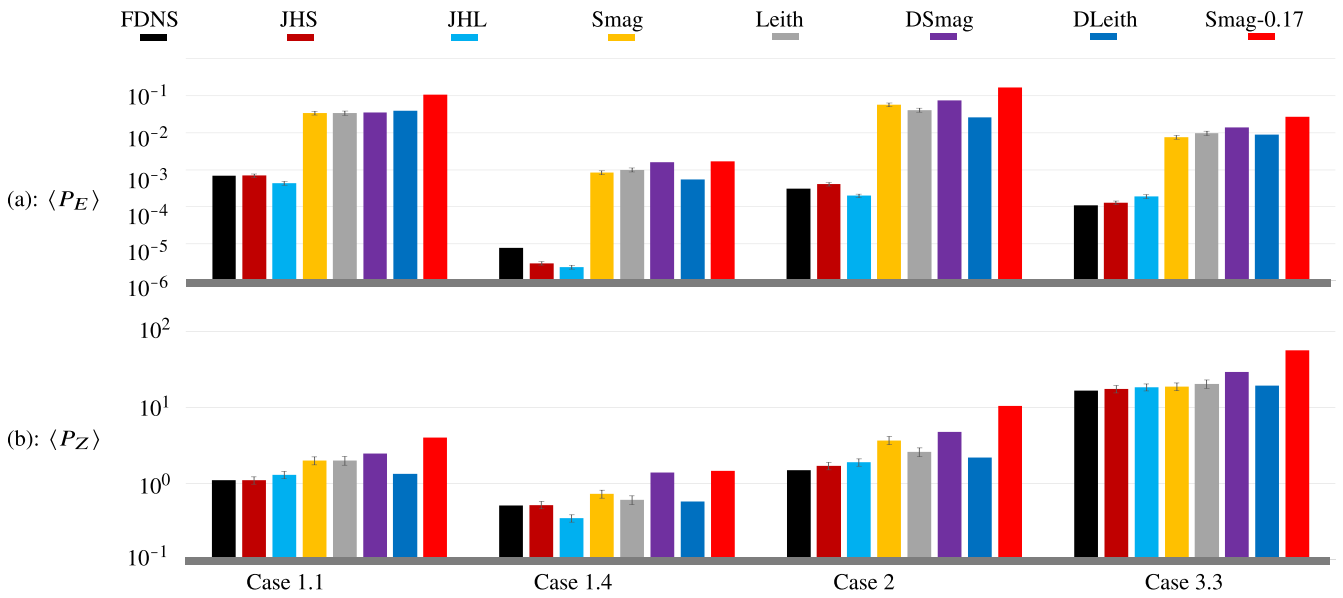

FIG. 3. Interscale transfers in representative cases calculated *a priori* for the same FDNS samples (note the logarithmic scale). EKI-optimized closures (JHS, JHL, Smag, and Leith) are compared against baselines (DSmag, DLeith, and Smag-0.17) and the truth (FDNS). The error bars on JHS, JHL, Smag, and Leith show the uncertainty in the interscale transfers due to the uncertainty of the EKI-optimized parameters. Values for all eight cases are shown in Tables I and II of the Supplemental Material [68].

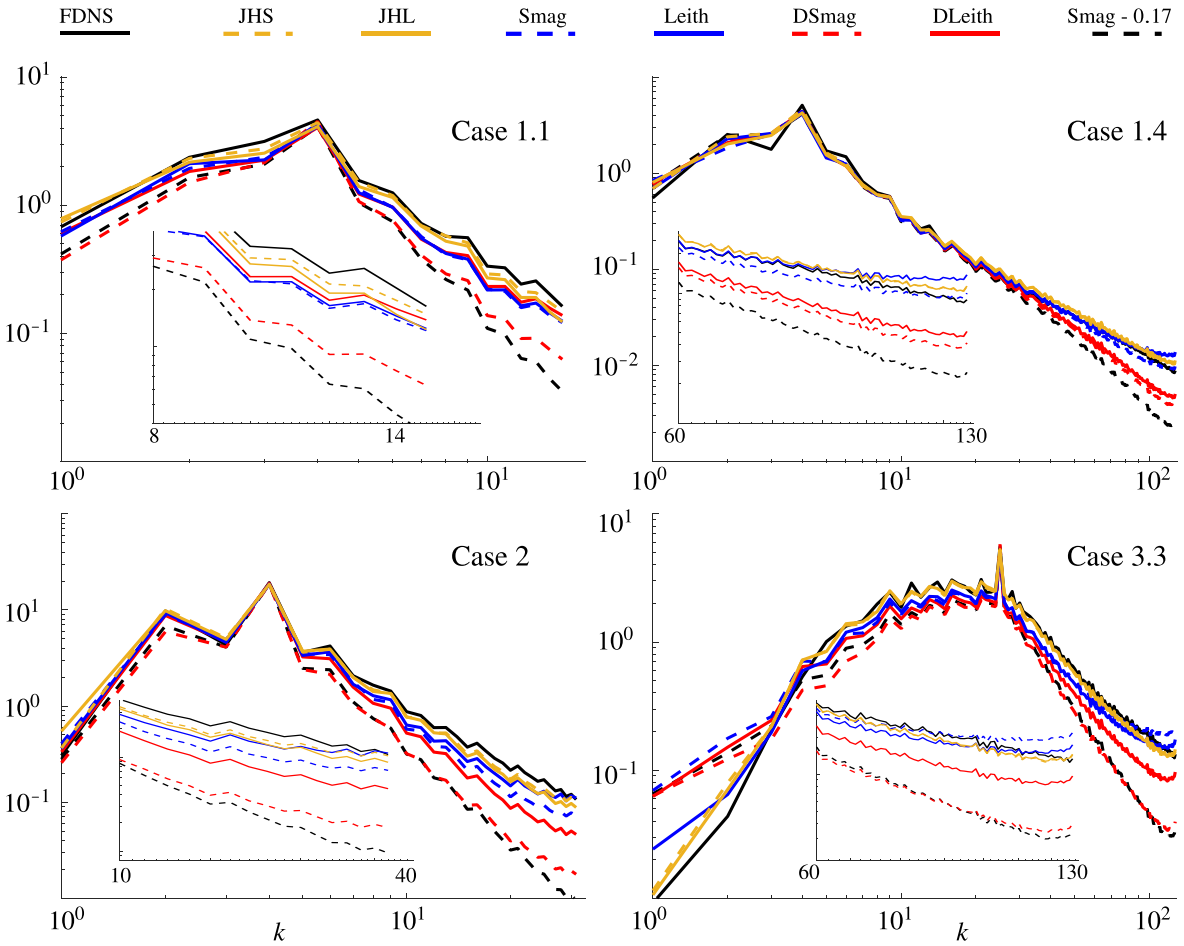

FIG. 4. Enstrophy spectra ($\hat{Z}$) from FDNS and LES for representative cases (see the Supplemental Material for other cases [68]). The insets magnify the high k part of the spectra.

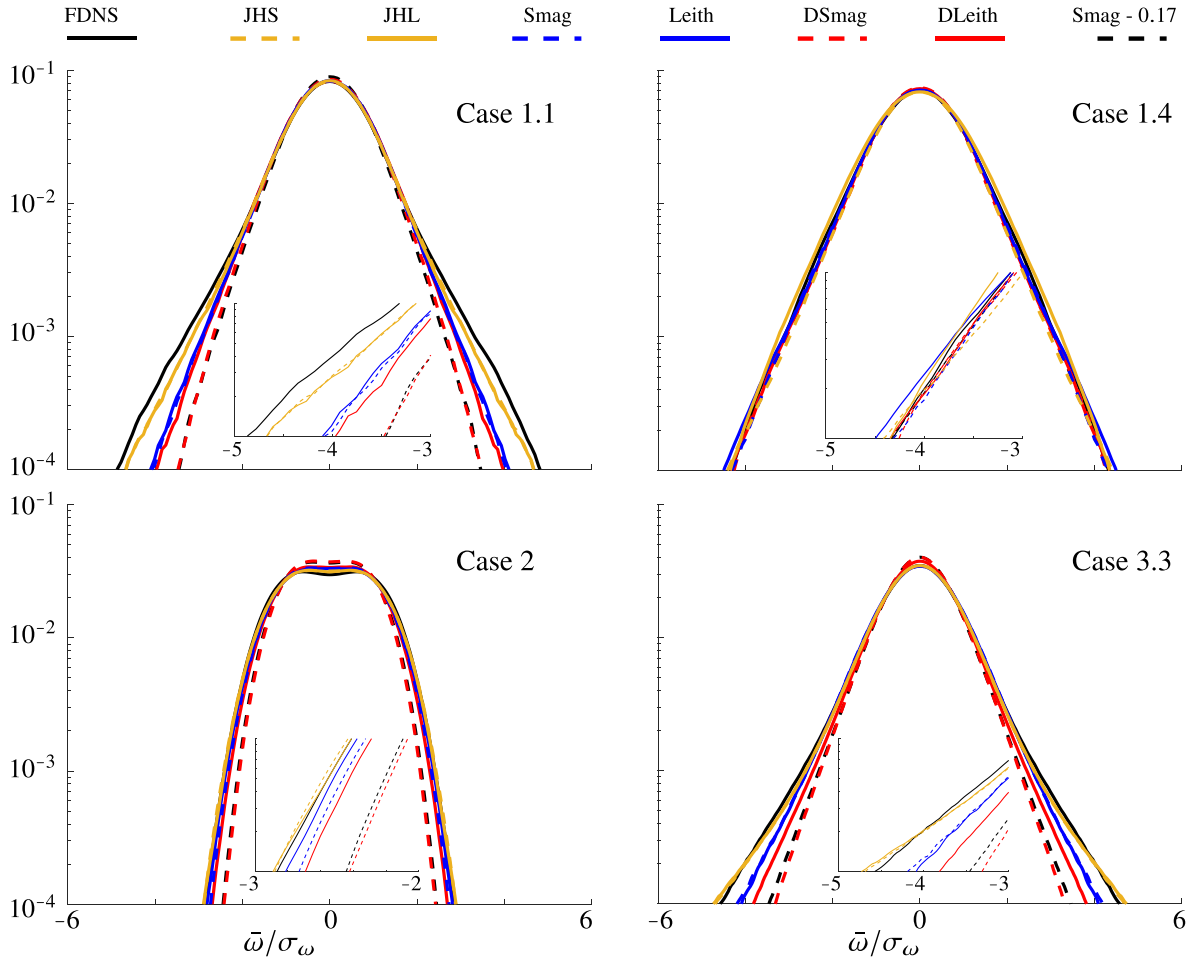

FIG. 5. Probability density functions of $\bar{\omega}$ normalized by the standard deviation of FDNS (σ_ω) for representative cases (see the Supplemental Material for other cases [68]). The insets close up into the left tails. A kernel estimator with a bandwidth of 1 is used to generate the probability density functions.

Figure 5 shows the probability density function $[\mathcal{P}(\bar{\omega})]$ for representative cases. Except for Cases 1.3 and 1.4 for which LES with all closures match the FDNS $\mathcal{P}(\bar{\omega})$ well (down to the tails), for all other cases, LES with optimized JH closures clearly outperform the baselines and optimized Smag and Leith in matching the FDNS $\mathcal{P}$, especially at the tails (rare, extreme events). In fact, except for Case 1.1, which involves the largest $N_{\mathrm{DNS}}/N_{\mathrm{LES}} = 32$, LES with optimized JH matches the FDNS $\mathcal{P}$ well. LES with optimized Smag and Leith also outperform the baselines (see, e.g., Sec. 3.3), although for Cases 1.1, 1.2, and 3.1, DLeith shows comparable performance.

B. Comparison with semianalytical estimates

To fully understand the EKI-optimized parameters, we compare the online-learned estimates with predictions from a semianalytical derivation that was recently introduced in Guan and Hassanzadeh [61]. The semianalytical derivation uses the interscale enstrophy transfer $\eta = \langle \bar{\omega}\Pi \rangle$ [24,86] and the TKE scaling law $\hat{E}(k) = A\eta^{2/3}k^{-3}$ [2,87,88] to obtain

$$C_{\mathrm{S}} = (A^3/2)^{-1/4}\pi^{-1}(\ln(k_c))^{-1/4}, \tag{6}$$

$$C_{\mathrm{L}} = 1/(\pi A^{1/2}), \tag{7}$$

and

$$C_{\mathrm{JHL}} = (A/2)^{-1/4}\pi^{-1}(1 - C_{\mathrm{B}}/\ln(k_c))^{-1/6}. \tag{8}$$

A in these equations can be obtained by linearly fitting the scaling law in the inertial region ($[k_f + 1, k_c]$) of the TKE spectra from a few DNS snapshots (see Table I for the estimated values). In the fitting process, η is obtained by averaging the value for the FDNS snapshots. The near universality of the online-learned C_{L}, C_{S}, and C_{JHL} can be explained by the consistency in $A \approx 1.8$–1.9 across most cases (except for Case 2) and the weak dependence on k_c [see Eqs. (6) and (8)]. The value of A for 2D homogeneous isotropic turbulence here is close to the one predicted using the renormalization-group analysis, i.e., $A = 1.923$ (e.g., Refs. [89,90]). In fact, for most 2D homogeneous isotropic turbulence, A is found empirically to be around 1.0–2.0 [57,91–94]. The difference in A for Case 2 is likely due to the β effect and the resulting strong anisotropy. Nevertheless, once A is determined, e.g., empirically using only the TKE spectra of a few DNS snapshots, the semianalytical derivation can be used to estimate the model parameters. Furthermore, the consistency between the EKI-optimized parameters and the semianalytical prediction within $\mathcal{O}(0.01)$ uncertainty suggests that these parameters are learned from data such that

they match the enstrophy interscale transfer (η) of LES to FDNS.

IV. SUMMARY AND CONCLUSION

We use the CES framework with EKI to learn *online* the parameters of four well-known physics-based models for seven setups of 2D homogeneous isotropic turbulence, plus one that includes a planetary vorticity gradient (β-plane effect), and interpret them with a new semianalytical derivation. As summarized below, we demonstrate that this approach is data efficient, generalizable, and interpretable, and further helps with gaining insight into the physics of turbulence modeling.

The objective of the learning process is to match the TKE spectrum $\hat{E}(k)$ of LES with the optimized closure with the spectrum of FDNS. The latter can be obtained from a short DNS run, making the approach *data efficient*. Note that we have also explored using the enstrophy spectrum $\hat{Z}(k)$ of FDNS as the target for learning C_S and C_L for Case 3.1, and found similar constant parameters (1.2 vs 1.3 and 2.4 vs 2.4), which is not surprising given the simple relationship between $\hat{Z}(k)$ and $\hat{E}(k)$. The estimated parameters of each closure are found to be nearly constant across the eight cases, suggesting that the optimized closures are *generalizable*. We further analyze these parameters using turbulence theory and scaling analysis and verify that the parameters are optimized to match the FDNS interscale enstrophy transfer. Such analysis provides further insight into these already *interpretable* closures. Furthermore, this comparison also provides further support for the semianalytical derivation, which involves several assumptions and simplifications. The semianalytical derivation has its own significance: To calculate the model parameters, e.g., C_S, C_L, and C_{JHL}, it only requires estimating A from a short high-resolution simulation, without any deep learning or online learning techniques.

A number of metrics are used to quantify the performance of the optimized closures. The *a priori* metrics involving mean interscale energy and enstrophy transfers (between the SGS and resolved flow) show major improvements resulting from optimization. In particular, the backscattering models (JHS and JHL) demonstrate significantly better ability in capturing these transfers compared to baselines and optimized eddy-viscosity models. *A posteriori* tests, focused on examining the enstrophy spectra at the largest and smallest length scales and the tails of the vorticity probability density functions, again show major improvements gained from the optimization. Overall, the optimized eddy-viscosity closures are found to work well (and better than the baselines) except for cases in which the LES resolution is low compared to the forcing length scale (i.e., small k_c/k_f). For such cases, the eddy-viscosity closures have structural errors, and optimization alone (which only addresses parametric errors) cannot further improve the LES. The JH closure, on the other hand, leads to LES that matches FDNS closely based on all metrics and for all cases, further showing the importance of combining theoretical advances and data-driven techniques to develop data-efficient frameworks for accurate, interpretable, and generalizable closures.

The next steps in this work should include exploring cases with more extreme regimes (e.g., higher Re) and different dynamics, and investigating more complex test cases (e.g., ocean general circulation models). The objectives will be to examine the performance and scalability of the framework, as well as gaining further physical insight through the semianalytical model. Furthermore, the applicability of the CES framework to using experimental or noisy data as target/reference should be explored. It is expected that the framework performs well as long as the key statistics (to be matched by the subgrid-scale model of LES) extracted from the data are largely invariant in time.

ACKNOWLEDGMENTS

We are grateful to two anonymous reviewers and the editor for insightful comments and suggestions. We thank Karan Jakhar, Malte Jansen, and Rambod Mojgani for insightful discussions. This work was supported by ONR Award No. N000142012722, NSF Grants No. OAC-2005123 and No. AGS-1835860, and by Schmidt Sciences, LLC. Computational resources were provided by ACCESS (ATM170020) and CISL (URIC0004).

DATA AVAILABILITY

The data that support the findings of this article are openly available [95,96]; embargo periods may apply.

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

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

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

[4] S. Khani and M. L. Waite, Backscatter in stratified turbulence, Eur. J. Mech. B Fluids **60**, 1 (2016).

[5] S. Juricke, S. Danilov, N. Koldunov, M. Oliver, D. Sein, D. Sidorenko, and Q. Wang, A kinematic kinetic energy backscatter parametrization: From implementation to global ocean simulations, J. Adv. Model. Earth Syst. **12**, e2020MS002175 (2020).

[6] I. Grooms, Backscatter in energetically-constrained Leith parameterizations, Ocean Modell. **186**, 102265 (2023).

[7] J. Bardina, J. Ferziger, and W. Reynolds, Improved subgrid-scale models for large-eddy simulation, in *13th Fluid and Plasmadynamics Conference* (AIAA, Snowmass, CO, 1980), p. 1357.

[8] M. F. Jansen and I. M. Held, Parameterizing subgrid-scale eddy effects using energetically consistent backscatter, Ocean Modell. **80**, 36 (2014).

[9] M. F. Jansen, I. M. Held, A. Adcroft, and R. Hallberg, Energy budget-based backscatter in an eddy permitting primitive equation model, Ocean Modell. **94**, 15 (2015).

[10] J. Berner, G. Shutts, M. Leutbecher, and T. Palmer, A spectral stochastic kinetic energy backscatter scheme and its impact on

flow-dependent predictability in the ECMWF ensemble prediction system, J. Atmos. Sci. **66**, 603 (2009).

[11] L. A. Mansfield and A. Sheshadri, Calibration and uncertainty quantification of a gravity wave parameterization: A case study of the quasi-biennial oscillation in an intermediate complexity climate model, J. Adv. Model. Earth Syst. **14**, e2022MS003245 (2022).

[12] T. Schneider, S. Behera, G. Boccaletti, C. Deser, K. Emanuel, R. Ferrari, L. R. Leung, N. Lin, T. Müller, A. Navarra, *et al.*, Harnessing AI and computing to advance climate modelling and prediction, Nat. Clim. Change **13**, 887 (2023).

[13] B. Sanderse, P. Stinis, R. Maulik, and S. E. Ahmed, Scientific machine learning for closure models in multiscale problems: A review, arXiv:2403.02913.

[14] C.-Y. Lai, P. Hassanzadeh, A. Sheshadri, M. Sonnewald, R. Ferrari, and V. Balaji, Machine learning for climate physics and simulations, Annu. Rev. Condens. Matter Phys. **16**, 343 (2025).

[15] V. Eyring, W. D. Collins, P. Gentine, E. A. Barnes, M. Barreiro, T. Beucler, M. Bocquet, C. S. Bretherton, H. M. Christensen, K. Dagon, *et al.*, Pushing the frontiers in climate modelling and analysis with machine learning, Nat. Climate Change **14** 916 (2024).

[16] A. Bracco, J. Brajard, H. A. Dijkstra, P. Hassanzadeh, C. Lessig, and C. Monteleoni, Machine learning for the physics of climate, Nat. Rev. Phys. **7**, 6 (2025).

[17] L. A. Mansfield and A. Sheshadri, Uncertainty quantification of a machine learning subgrid-scale parameterization for atmospheric gravity waves, J. Adv. Model. Earth Syst. **16**, e2024MS004292 (2024).

[18] R. Maulik, O. San, A. Rasheed, and P. Vedula, Subgrid modelling for two-dimensional turbulence using neural networks, J. Fluid Mech. **858**, 122 (2019).

[19] A. Beck, D. Flad, and C.-D. Munz, Deep neural networks for data-driven LES closure models, J. Comput. Phys. **398**, 108910 (2019).

[20] L. Zanna and T. Bolton, Data-driven equation discovery of ocean mesoscale closures, Geophys. Res. Lett. **47**, e2020GL088376 (2020).

[21] Y. Guan, A. Chattopadhyay, A. Subel, and P. Hassanzadeh, Stable a posteriori LES of 2D turbulence using convolutional neural networks: Backscattering analysis and generalization to higher Re via transfer learning, J. Comput. Phys. **458**, 111090 (2022).

[22] K. Jakhar, Y. Guan, R. Mojgani, A. Chattopadhyay, and P. Hassanzadeh, Learning closed-form equations for subgrid-scale closures from high-fidelity data: Promises and challenges, J. Adv. Model. Earth Syst. **16**, e2023MS003874 (2024).

[23] K. Jakhar, Y. Guan, and P. Hassanzadeh, Analytical and AI-discovered stable, accurate, and generalizable subgrid-scale closure for geophysical turbulence, Phys. Rev. Lett. **136**, 064201 (2026).

[24] Y. Guan, A. Subel, A. Chattopadhyay, and P. Hassanzadeh, Learning physics-constrained subgrid-scale closures in the small-data regime for stable and accurate LES, Physica D **443**, 133568 (2023).

[25] K. Srinivasan, M. D. Chekroun, and J. C. McWilliams, Turbulence closure with small, local neural networks: Forced two-dimensional and β-plane flows, J. Adv. Model. Earth Syst. **16**, e2023MS003795 (2024).

[26] A. Ross, Z. Li, P. Perezhogin, C. Fernandez-Granda, and L. Zanna, Benchmarking of machine learning ocean subgrid parameterizations in an idealized model, J. Adv. Model. Earth Syst. **15**, e2022MS003258 (2023).

[27] J. Yuval and P. A. O'Gorman, Neural-network parameterization of subgrid momentum transport in the atmosphere, J. Adv. Model. Earth Syst. **15**, e2023MS003606 (2023).

[28] S. Rasp, M. S. Pritchard, and P. Gentine, Deep learning to represent subgrid processes in climate models, Proc. Natl. Acad. Sci. USA **115**, 9684 (2018).

[29] A. Subel, Y. Guan, A. Chattopadhyay, and P. Hassanzadeh, Explaining the physics of transfer learning in data-driven turbulence modeling, PNAS Nexus **2**, pgad015 (2023).

[30] T. Beucler, P. Gentine, J. Yuval, A. Gupta, L. Peng, J. Lin, S. Yu, S. Rasp, F. Ahmed, P. A. O'Gorman, *et al.*, Climate-invariant machine learning, Sci. Adv. **10**, eadj7250 (2024).

[31] Y. Q. Sun, P. Hassanzadeh, M. J. Alexander, and C. G. Kruse, Quantifying 3D gravity wave drag in a library of tropical convection-permitting simulations for data-driven parameterizations, J. Adv. Model. Earth Syst. **15**, e2022MS003585 (2023).

[32] C. Christopoulos, I. Lopez-Gomez, T. Beucler, Y. Cohen, C. Kawczynski, O. R. Dunbar, and T. Schneider, Online learning of entrainment closures in a hybrid machine learning parameterization, J. Adv. Model. Earth Syst. **16**, e2024MS004485 (2024).

[33] T. Schneider, A. M. Stuart, and J.-L. Wu, Ensemble Kalman inversion for sparse learning of dynamical systems from time-averaged data, J. Comput. Phys. **470**, 111559 (2022).

[34] M. C. Kennedy and A. O'Hagan, Bayesian calibration of computer models, J. R. Stat. Soc. B **63**, 425 (2001).

[35] J.-L. Wu, M. E. Levine, T. Schneider, and A. Stuart, Learning about structural errors in models of complex dynamical systems, J. Comput. Phys. **513** 113157 (2024).

[36] H. A. Pahlavan, P. Hassanzadeh, and M. J. Alexander, Explainable offline-online training of neural networks for parameterizations: A 1D gravity wave-QBO testbed in the small-data regime, Geophys. Res. Lett. **51**, e2023GL106324 (2024).

[37] G. Novati, H. L. de Laroussilhe, and P. Koumoutsakos, Automating turbulence modelling by multi-agent reinforcement learning, Nat. Mach. Intell. **3**, 87 (2021).

[38] Y. Guan, L. Amoudruz, S. Litvinov, K. Jakhar, R. Mojgani, P. Koumoutsakos, and P. Hassanzadeh, Prediction of extreme events in multiscale simulations of geophysical turbulence using reinforcement learning, arXiv:2603.03351.

[39] M. Buzzicotti and P. Clark Di Leoni, Synchronizing subgrid scale models of turbulence to data, Phys. Fluids **32**, 125116 (2020).

[40] 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).

[41] J. Sirignano, J. F. MacArt, and J. B. Freund, DPM: A deep learning PDE augmentation method with application to large-eddy simulation, J. Comput. Phys. **423**, 109811 (2020).

[42] H. Frezat, J. Le Sommer, R. Fablet, G. Balarac, and R. Lguensat, A posteriori learning for quasi-geostrophic turbulence parametrization, J. Adv. Model. Earth Syst. **14**, e2022MS003124 (2022).

[43] B. List, L.-W. Chen, and N. Thuerey, Learned turbulence modelling with differentiable fluid solvers: Physics-based loss functions and optimisation horizons, J. Fluid Mech. **949**, A25 (2022).

[44] V. Shankar, D. Chakraborty, V. Viswanathan, and R. Maulik, Differentiable turbulence: Closure as a partial differential equation constrained optimization, Phys. Rev. Fluids **10**, 024605 (2025).

[45] T. Schneider, N. Jeevanjee, and R. Socolow, Accelerating progress in climate science, Phys. Today **74**(6), 44 (2021).

[46] T. Schneider, L. R. Leung, and R. C. Wills, Opinion: Optimizing climate models with process-knowledge, resolution, and Artificial Intelligence, Atmospheric Chem. Phys. **24**, 7041 (2024).

[47] I. Lopez-Gomez, C. Christopoulos, H. L. Langeland Ervik, O. R. Dunbar, Y. Cohen, and T. Schneider, Training physics-based machine-learning parameterizations with gradient-free ensemble Kalman methods, J. Adv. Model. Earth Syst. **14**, e2022MS003105 (2022).

[48] X.-L. Zhang, H. Xiao, X. Luo, and G. He, Ensemble Kalman method for learning turbulence models from indirect observation data, J. Fluid Mech. **949**, A26 (2022).

[49] V. R. Martinez, J. Murri, and J. P. Whitehead, Relaxation-based schemes for on-the-fly parameter estimation in dissipative dynamical systems, Inverse Probl. **41**, 055001 (2025).

[50] E. Cleary, A. Garbuno-Inigo, S. Lan, T. Schneider, and A. M. Stuart, Calibrate, emulate, sample, J. Comput. Phys. **424**, 109716 (2021).

[51] M. Iglesias and Y. Yang, Adaptive regularisation for ensemble Kalman inversion, Inverse Probl. **37**, 025008 (2021).

[52] S. Vernon, E. Bach, and O. R. Dunbar, Nesterov acceleration for ensemble Kalman inversion and variants, J. Comput. Phys. **535** 114063 (2025).

[53] O. R. Dunbar, A. Garbuno-Inigo, T. Schneider, and A. M. Stuart, Calibration and uncertainty quantification of convective parameters in an idealized GCM, J. Adv. Model. Earth Syst. **13**, e2020MS002454 (2021).

[54] M. F. Howland, O. R. A. Dunbar, and T. Schneider, Parameter uncertainty quantification in an idealized GCM with a seasonal cycle, J. Adv. Model. Earth Syst. **14**, e2021MS002735 (2022).

[55] A. N. Souza, G. L. Wagner, A. Ramadhan, B. Allen, V. Churavy, J. Schloss, J. Campin, C. Hill, A. Edelman, J. Marshall, G. Flierl, and R. Ferrari, Uncertainty quantification of ocean parameterizations: Application to the k-profile-parameterization for penetrative convection, J. Adv. Model. Earth Syst. **12**, e2020MS002108 (2020).

[56] M. Bieli, O. R. A. Dunbar, E. K. de Jong, A. Jaruga, T. Schneider, and T. Bischoff, An efficient bayesian approach to learning droplet collision kernels: Proof of concept using "Cloudy," a new *n*-moment bulk microphysics scheme, J. Adv. Model. Earth Syst. **14**, e2022MS002994 (2022).

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

[58] P. A. Davidson, *Turbulence: An Introduction for Scientists and Engineers* (Oxford University Press, Oxford, UK, 2015).

[59] G. K. Vallis, *Atmospheric and Oceanic Fluid Dynamics* (Cambridge University Press, Cambridge, UK, 2017).

[60] B. Gallet and R. Ferrari, The vortex gas scaling regime of baroclinic turbulence, Proc. Natl. Acad. Sci. USA **117**, 4491 (2020).

[61] Y. Guan and P. Hassanzadeh, Semi-analytical eddy-viscosity and backscattering closures for 2D geophysical turbulence, arXiv:2504.09670.

[62] P. Sagaut, *Large Eddy Simulation for Incompressible Flows* (Springer Science & Business Media, Berlin, Germany, 2006).

[63] D. Lilly, The representation of small-scale turbulence in numerical simulation experiments, NCAR Report, National Center For Atmospheric Research, Boulder, Colorado, 1967.

[64] S. B. Pope, *Turbulent Flows* (IOP Publishing, Cambridge, UK, 2001).

[65] O. J. McMillan and J. H. Ferziger, Direct testing of subgrid-scale models, AIAA J. **17**, 1340 (1979).

[66] M. Germano, U. Piomelli, P. Moin, and W. H. Cabot, A dynamic subgrid-scale eddy viscosity model, Phys. Fluids **3**, 1760 (1991).

[67] S. Khani and M. L. Waite, Large eddy simulations of stratified turbulence: The dynamic Smagorinsky model, J. Fluid Mech. **773**, 327 (2015).

[68] See Supplemental Material at http://link.aps.org/supplemental/10.1103/mnbm-3g56 for additional details on the DNS setup, FDNS training data, EKI algorithm and convergence, interscale transfer statistics, computational costs, and supplemental results for additional cases.

[69] P. Chandramouli, E. Memin, and D. Heitz, 4D large scale variational data assimilation of a turbulent flow with a dynamics error model, J. Comput. Phys. **412**, 109446 (2020).

[70] M. A. Iglesias, K. J. Law, and A. M. Stuart, Ensemble Kalman methods for inverse problems, Inverse Probl. **29**, 045001 (2013).

[71] Y. Chen and D. S. Oliver, Ensemble randomized maximum likelihood method as an iterative ensemble smoother, Math. Geosci. **44**, 1 (2012).

[72] M. Bocquet and P. Sakov, An iterative ensemble Kalman smoother, Q. J. R. Meteorol. Soc. **140**, 1521 (2014).

[73] V. Mons, Y. Du, and T. A. Zaki, Ensemble-variational assimilation of statistical data in large-eddy simulation, Phys. Rev. Fluids **6**, 104607 (2021).

[74] D. A. Buchta and T. A. Zaki, Observation-infused simulations of high-speed boundary-layer transition, J. Fluid Mech. **916**, A44 (2021).

[75] T. A. Zaki, Turbulence from an observer perspective, Annu. Rev. Fluid Mech. **57**, 311 (2024).

[76] C. Liu, Q. Xiao, and B. Wang, An ensemble-based four-dimensional variational data assimilation scheme. Part I: Technical formulation and preliminary test, Mon. Weather Rev. **136**, 3363 (2008).

[77] J. Bergstra and Y. Bengio, Random search for hyper-parameter optimization, J. Mach. Learn. Res. **13**, 281 (2012).

[78] V. Tomar, M. Bansal, and P. Singh, Metaheuristic algorithms for optimization: A brief review, Eng. Proc. **59**, 238 (2024).

[79] S. Chen, J. Montgomery, and A. Bolufé-Röhler, Measuring the curse of dimensionality and its effects on particle swarm optimization and differential evolution, Appl. Intell. **42**, 514 (2015).

[80] E. Calvello, S. Reich, and A. M. Stuart, Ensemble Kalman methods: A mean-field perspective, Acta Numer. **34**, 123 (2025).

[81] O. R. a. Dunbar, I. Lopez-Gomez, A. Garbuno-Iñigo, D. Z. Huang, E. Bach, and J.-l. Wu, Ensemblekalmanprocesses.jl: Derivative-free ensemble-based model calibration, J. Open Source Software **7**, 4869 (2022).

[82] A. Garbuno-Inigo, F. Hoffmann, W. Li, and A. M. Stuart, Interacting Langevin diffusions: Gradient structure and ensemble Kalman sampler, SIAM J. Appl. Dyn. Syst. **19**, 412 (2020).

[83] A. Garbuno-Inigo, N. Nüsken, and S. Reich, Affine invariant interacting Langevin dynamics for Bayesian inference, SIAM J. Appl. Dyn. Syst. **19**, 1633 (2020).

[84] O. R. Dunbar, M. Bieli, A. Garbuno-Iñigo, M. Howland, A. N. de Souza, L. A. Mansfield, G. L. Wagner, N. Efrat-Henrici, *et al.*, Calibrateemulatesample. jl: Accelerated parametric uncertainty quantification, J. Open Source Software **9**, 6372 (2024).

[85] R. Maulik and O. San, Dynamic modeling of the horizontal eddy viscosity coefficient for quasigeostrophic ocean circulation problems, J. Ocean Eng. Sci. **1**, 300 (2016).

[86] J. Thuburn, J. Kent, and N. Wood, Cascades, backscatter and conservation in numerical models of two-dimensional turbulence, Q. J. R. Meteorol. Soc. **140**, 626 (2014).

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

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

[89] P. Olla, Renormalization-group analysis of two-dimensional incompressible turbulence, Phys. Rev. Lett. **67**, 2465 (1991).

[90] M. K. Nandy and J. K. Bhattacharjee, Mode-coupling theory, dynamic scaling, and two-dimensional turbulence, Int. J. Mod. Phys. B **09**, 1081 (1995).

[91] L. M. Smith and V. Yakhot, Bose condensation and small-scale structure generation in a random force driven 2D turbulence, Phys. Rev. Lett. **71**, 352 (1993).

[92] T. Gotoh, Energy spectrum in the inertial and dissipation ranges of two-dimensional steady turbulence, Phys. Rev. E **57**, 2984 (1998).

[93] E. Lindborg and A. Vallgren, Testing Batchelor's similarity hypotheses for decaying two-dimensional turbulence, Phys. Fluids **22**, 091704 (2010).

[94] A. Gupta, R. Jayaram, A. G. Chaterjee, S. Sadhukhan, R. Samtaney, and M. K. Verma, Energy and enstrophy spectra and fluxes for the inertial-dissipation range of two-dimensional turbulence, Phys. Rev. E **100**, 053101 (2019).

[95] Y. Guan and P. Hassanzadeh, Code and data for 2D turbulence DNS and LES solver, GitHub (2025), https://github.com/envfluids/2D-DDP.

[96] CliMA, Code and data for calibrate, emulate, and sample (CES), GitHub (2025), https://clima.github.io/CalibrateEmulateSample.jl/dev/.