
Article:
Li, J., Tian, M. and Li, Y. orcid.org/0000-0001-7907-5176 (2022) Synchronizing large eddy simulations with direct numerical simulations via data assimilation. Physics of Fluids, 34 (6). 065108. ISSN 1070-6631

https://doi.org/10.1063/5.0089895

Physics of Fluids

RESEARCH ARTICLE | JUNE 02 2022

Synchronizing large eddy simulations with direct numerical simulations via data assimilation

Jian Li; Mengdan Tian; Yi Li

Physics of Fluids 34, 065108 (2022)
https://doi.org/10.1063/5.0089895

Physics of Fluids ARTICLE scitation.org/journal/phf

Synchronizing large eddy simulations with direct numerical simulations via data assimilation

Cite as: Phys. Fluids **34**, 065108 (2022); doi: 10.1063/5.0089895
Submitted: 1 March 2022 · Accepted: 13 May 2022 ·
Published Online: 2 June 2022

Jian Li,[1] Mengdan Tian,[1] and Yi Li[2,a)]

AFFILIATIONS

[1]School of Naval Architecture and Maritime, Zhejiang Ocean University, Zhoushan 316022, China
[2]School of Mathematics and Statistics, University of Sheffield, Sheffield S3 7RH, United Kingdom

[a)]Author to whom correspondence should be addressed: yili@sheffield.ac.uk

ABSTRACT

The synchronization of large eddy simulations to direct numerical simulations via a data assimilation scheme is investigated in Kolmogorov flows, where the large scales of the velocity field in large eddy simulations are replaced by those in the direct numerical simulations. We show that, when the amount of assimilated data exceeds a threshold given by a threshold wavenumber, all large eddy simulations with the same subgrid-scale model converge to an orbit that is synchronized with the direct numerical simulations in phase. The threshold wavenumbers for the standard and dynamic Smagorinsky models are smaller than those for the dynamic mixed model and are reduced when the filter scale increases. The error in the synchronized large eddy simulations is examined in detail. We reveal that for larger filter scales, unexpectedly, the velocity simulated with the standard and the dynamic Smagorinsky models can be more accurate than the one calculated with the dynamic mixed model. The robustness of the results is assessed in simulations where the assimilated data are perturbed by random noise and in homogeneous turbulence which is driven by a linear forcing term. Good synchronization is still obtained in both cases. The Smagorinsky models still display better performance than the dynamic mixed model.

I. INTRODUCTION

Data assimilation (DA) is a practice where one synthesizes observational data and numerical simulations to improve prediction. The approach has been used in numerical weather prediction, but it has only recently been applied to the simulations of turbulent flows. A review of various DA approaches can be found in Kalnay.[1] The common methods being used in turbulence simulations include the variational approaches,[2–9] the ensemble approaches,[1,10] and other sequential assimilation methods.[11–17] In many cases, DA is used to improve the simulations based on the Reynolds-averaged Navier–Stokes equation (NSE),[3,9,10,18] but it has recently been coupled with large eddy simulations (LES) or direct numerical simulations (DNS). The aim, in this case, is to recover the chaotic instantaneous turbulent fields based on incomplete measurement data. Thus, the problem can be viewed as one of chaos synchronization.[19,20] The most simple form of synchronization is the so-called complete synchronization, where the chaotic orbits of two identical dynamical systems exactly replicate each other asymptotically. Yoshida, Yamaguchi, and Kaneda[11] investigate the complete synchronization of two isotropic turbulent flows via a simple data assimilation scheme where the Fourier modes from one flow with wavenumbers less than k_m are replaced by those in the other. It is found that complete synchronization is achieved only if $k_m > k_c$, where $k_c\eta \approx 0.2$ with η being the Kolmogorov length scale. The problem is revisited in Lalescu, Meneveau, and Eyink[12] with a different forcing scheme as well as anisotropic grids, and $k_c\eta \approx 0.15$ is found. When k_m is smaller than k_c, only partial synchronization can be achieved. Vela-Martin[15] documents the features of partial synchronization in isotropic turbulence and shows that strong vorticity is better synchronized than weaker vorticity. Nikolaidis and Ioannou[16] show that Couette flows can be synchronized by assimilating streamwise Fourier modes with wavelength exceeding a threshold value. Channel flows are investigated by Wang and Zaki.[17] Data from layers with different orientations are assimilated. Scaling of the thickness of the layers needed for synchronization is established. Nudging is used in Leoni, Mazzino, and Biferale[13,14] to synchronize isotropic turbulence with or without rotation and infer system parameters and reconstruct the large-scale structures of the flows.

These recent works focus on the synchronization between two DNS with identical system parameters. In reality, the simulations of high Reynolds number flows often rely on coarse-grained models such as LES. Thus, one needs to ask what is achievable when one synchronizes LES with the data obtained from the full flow. This problem is addressed in Buzzicotti and Leoni[21] using the nudging method, in which the filtered Navier–Stokes equation (NSE) is supplemented by a linear forcing term which nudges the LES velocity toward the

DNS one. Nudging is applied at all length scales with fairly strong nudging strength. It is found that the error between the nudged LES velocity and DNS velocity can be minimized by choosing the parameters in the subgrid-scale (SGS) models judiciously. The observation shows that nudging (and more generally DA) can be used to optimize model parameters. More recently, Zauner *et al.*[22] apply nudging in a simulation of the flow around a square cylinder. They observe the synchronization of low-frequency vortex shedding with relatively sparse data although the model they used is the unsteady Reynolds-averaged Navier–Stokes equation rather than LES.

Despite these recent works, several questions remain unclear regarding the synchronization between LES and DNS. First of all, because LES and DNS are non-identical systems, only partial synchronization in the form of phase synchronization (PS) can be achieved, where the phases of the two systems converge to each other asymptotically while the amplitudes remain out of sync.[20] One key question regarding this partial synchronization is how much data are needed to achieve it for different subgrid-scale (SGS) models because arguably an SGS model requiring less data to sync with the DNS potentially is advantageous. Our first objective is to answer this question. The second objective is to quantify and compare the error of the partially synchronized LES velocity for several canonical SGS models, with the synchronization of the vorticity given specific attention. We consider turbulent flows in a three-dimensional (3D) periodic box and focus on the DA scheme used in Yoshida, Yamaguchi, and Kaneda.[11] Three SGS models, the standard Smagorinsky model (SSM), the dynamic Smagorinsky model (DSM), and the dynamic mixed model (DMM), are compared. To obtain information on the robustness of the results, we also consider cases where the assimilated data are contaminated by noise and a different flow which is driven by a different forcing mechanism.

We introduce the governing equations and the methods in Sec. II. The numerical methods and the results are discussed in Sec. III. Section IV summarizes the main observations from the results.

II. GOVERNING EQUATIONS

To demonstrate phase synchronization, traditionally, one would define the phase of the systems and show that the phase synchronizes over time. However, there is no natural way to define the phase of a 3D turbulent velocity field. We, thus, use the following indirect approach. We consider a master system M and two identical slave systems A and B that are different from M (c.f. Lalescu, Meneveau, and Eyink[12]). The data from M are assimilated into A and B in exactly the same way. Since M and A are governed by different equations, it is impossible to achieve complete synchronization between them (or between M and B). However, if complete synchronization is achieved between A and B, then their states converge toward an orbit which depends only on the assimilated data from M. We may then assert that this orbit of A and B is in phase synchronization with the orbit of M.

Therefore, in this investigation, we consider a master system M which is governed by the NSE, and two identical slave systems A and B which are governed by the filtered NSE (fNSE). The NSE reads

$$\partial_t \boldsymbol{u} + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} = -\nabla p + \nu\nabla^2\boldsymbol{u} + \boldsymbol{f}, \tag{1}$$

where $\boldsymbol{u}$ is the velocity, p is the pressure (divided by the constant density), ν is the viscosity, and $\boldsymbol{f}$ is the forcing term. In the majority of cases, we will consider

$$\boldsymbol{f} \equiv (a_f \cos k_f x_2, 0, 0) \tag{2}$$

with $a_f = 0.15$ and $k_f = 1$. The velocity is assumed to be incompressible so that

$$\nabla\cdot\boldsymbol{u} = 0. \tag{3}$$

Customarily, the flow driven by forcing terms of this type is called the Kolmogorov flow.

The filtered Navier–Stokes equation (fNSE) is given by

$$\partial_t\overline{\boldsymbol{u}} + (\overline{\boldsymbol{u}}\cdot\nabla)\overline{\boldsymbol{u}} = -\nabla\overline{p} + \nabla\cdot(-\boldsymbol{\tau}) + \nu\nabla^2\overline{\boldsymbol{u}} + \overline{\boldsymbol{f}} + \boldsymbol{f}_s, \tag{4}$$

where the over-line $\bar{\ }$ represents filtering with filter length Δ, $\overline{\boldsymbol{u}}$ is the filtered velocity, $\overline{p}$ is the filtered pressure, and $\boldsymbol{\tau}$ is the SGS stress tensor with

$$\tau_{ij} = \overline{u_i u_j} - \overline{u}_i\overline{u}_j.$$

When multiple slave systems are simulated together with the master system, the same SGS model is used in the slave systems, so they are identical except for the initial velocity.

The additional forcing term $\boldsymbol{f}_s$ in Eq. (4) symbolically represents the effects of DA. The DNS (i.e., the master system M) and the LES (the slave systems) will be simulated concurrently. At each time step, the Fourier modes of the LES velocity field with $|\boldsymbol{k}| \le k_m$, $\boldsymbol{k}$ being the wavenumber, are replaced by those from the DNS velocity field. As such, these Fourier modes in the DNS are continuously assimilated into the LES velocity fields.

We compare three canonical SGS models. These models have been extensively studied; therefore, only the necessary formulae are given. For details see, e.g., Meneveau and Katz.[23] The first model is the standard Smagorinsky model (SSM) with

$$\tau_{ij} = -2(c_s\Delta)^2|\overline{\boldsymbol{s}}|\overline{s}_{ij}, \tag{5}$$

where $c_s = 0.16$ is the Smagorinsky coefficient, Δ is the filter length, and $\overline{s}_{ij}$ is the filtered strain rate tensor given by

$$\overline{s}_{ij} = \frac{1}{2}(\partial_j\overline{u}_i + \partial_i\overline{u}_j) \tag{6}$$

and $|\overline{\boldsymbol{s}}| \equiv (2\overline{s}_{ij}\overline{s}_{ij})^{1/2}$. The second model is the dynamic Smagorinsky model (DSM) which is also defined by Eq. (5), but c_s^2 is calculated dynamically from

$$c_s^2 = \frac{\langle L_{ij}M_{ij}\rangle_v}{\langle M_{ij}M_{ij}\rangle_v}, \tag{7}$$

where $\langle\ \rangle_v$ represents volume average, and

$$M_{ij} = -2\Delta^2\left[4|\widetilde{\overline{\boldsymbol{s}}}|\widetilde{\overline{s}}_{ij} - \widetilde{|\overline{\boldsymbol{s}}|\overline{s}_{ij}}\right], \quad L_{ij} = \widetilde{\overline{u}_i\overline{u}_j} - \widetilde{\overline{u}}_i\widetilde{\overline{u}}_j, \tag{8}$$

in which $\tilde{\ }$ denotes filtering with filter scale 2Δ. Equation (7) is derived from the Germano identity following the dynamic procedure.[23] Finally, the dynamic mixed model (DMM) is also considered, which is defined by

$$\tau_{ij} = -2(c_s\Delta)^2|\overline{\boldsymbol{s}}|\overline{s}_{ij} + c_{nl}\Delta^2\partial_k\overline{u}_i\partial_k\overline{u}_j. \tag{9}$$

Following the dynamic procedure, the equations for c_s^2 and c_{nl} are found to be

TABLE I. Parameters for the DNS. N_{DNS}^3 (N_{LES}^3): number of grid points in DNS (LES). δt: time step size used in both DNS and LES. δx_{DNS}: the grid size in DNS. u': root-mean-square velocity. ν: viscosity, ϵ: average viscous energy dissipation rate, λ: Taylor length scale, $Re_\lambda \equiv u'\lambda/\nu$: the Taylor–Reynolds number, and $\eta \equiv (\nu^3/\epsilon)^{1/4}$: Kolmogorov length scale. $\tau_K \equiv (\nu/\epsilon)^{1/2}$: Kolmogorov length scale.

N_{DNS}	Re_λ	N_{LES}	$\delta t u'/\delta x_{\mathrm{DNS}}$	u'	ϵ	η	ν	λ	τ_K
128	75	64, 96	0.074	0.63	0.072	0.042	0.0060	0.71	0.30
192	90	64, 96, 128	0.077	0.65	0.074	0.033	0.0044	0.61	0.24
256	112	64, 96, 128	0.078	0.66	0.077	0.024	0.0020	0.51	0.20

$$c_s^2 = \frac{\langle L_{ij}M_{ij}\rangle_v\langle N_{ij}N_{ij}\rangle_v - \langle L_{ij}N_{ij}\rangle_v\langle M_{ij}N_{ij}\rangle_v}{\langle M_{ij}M_{ij}\rangle_v\langle N_{ij}N_{ij}\rangle_v - \langle M_{ij}N_{ij}\rangle_v^2}, \tag{10}$$

$$c_{nl} = \frac{\langle L_{ij}N_{ij}\rangle_v\langle M_{ij}M_{ij}\rangle_v - \langle L_{ij}M_{ij}\rangle_v\langle M_{ij}N_{ij}\rangle_v}{\langle M_{ij}M_{ij}\rangle_v\langle N_{ij}N_{ij}\rangle_v - \langle M_{ij}N_{ij}\rangle_v^2}, \tag{11}$$

where

$$N_{ij} = \Delta^2\left[4\widetilde{\overline{A}}_{ik}\widetilde{\overline{A}}_{jk} - \widetilde{\overline{A}_{ik}\overline{A}_{jk}}\right] \tag{12}$$

with $\overline{A}_{ij} \equiv \partial_j\overline{u}_i$ being the filtered velocity gradient.

III. NUMERICAL SIMULATIONS AND RESULTS

Equations (1) and (4) are solved with the pseudo-spectral method in a periodic box $V = [0, 2\pi]^3$ which is discretized uniformly with N_{DNS} and N_{LES} grid points, respectively, in each direction for the DNS and the LES. The two-thirds rule[24] is used to dealiase the advection term so that the maximum effective wavenumber is $N_{\mathrm{DNS}}/3$ for the DNS and $N_{\mathrm{LES}}/3$ for the LES. Time stepping uses an explicit second-order Euler scheme with a first-order predictor and a second-order corrector based on the trapezoid rule.[6]

The Fourier components of the initial velocity field are randomly sampled from a Gaussian distribution with unit variance and then are rescaled so that the energy spectrum $E(k)$ of the velocity field is given by

$$E(k) = \frac{16\sqrt{2}u_0^2}{\sqrt{\pi}}\frac{k^4}{k_p^5}e^{-2k^2/k_p^2}, \tag{13}$$

where $u_0 = 0.66$ and $k_p = 3$. As such, the total kinetic energy of the velocity field is $3u_0^2/2$. The energy spectrum peaks at k_p and falls off rapidly as k increases so that practically the initial field contains only large-scale motions.

Starting from the initial velocity field, the simulations are run for a period of time until the flows are fully developed in which state the statistics of the velocity becomes stationary. At this point, data assimilation is initiated. As a result, the results presented below do not depend on the initialization of the simulations.

Table I summarizes the parameters for the DNS from which the assimilated data are extracted and the number of grid points for the corresponding LES. To investigate how the amount of data being assimilated affects the results, for each set of parameters in the table, multiple values of k_m are computed. In total, around 1100 runs of DNS and LES are conducted.

To fully resolve the small scales of the flows in DNS, it requires $k_{\max}\eta \geq 1.5$, where $k_{\max}$ is the maximum effective wavenumber.[24] In the current investigation, $k_{\max}\eta$ is 1.8, 2.1, and 2.0 for the cases with N_{DNS} being 128, 192, and 256, respectively. Therefore, the flows are well resolved.

A. The phase synchronization between LES and DNS

We begin with the first objective: the complete synchronization of two identical LES (slave systems A and B, with different initial conditions) with data assimilated from the same DNS (system M). The three systems are simulated concurrently.

The synchronization between A and B is measured with the synchronization error

$$\mathcal{E}_{AB}(t) = \frac{1}{2(2\pi)^3}\int_V (\overline{\boldsymbol{u}}^A - \overline{\boldsymbol{u}}^B)\cdot(\overline{\boldsymbol{u}}^A - \overline{\boldsymbol{u}}^B)dV, \tag{14}$$

with $\overline{\boldsymbol{u}}^A$ and $\overline{\boldsymbol{u}}^B$ being the velocity in systems A and B, respectively. The results for selected cases are shown in Fig. 1 to illustrate the general features. The lines represent $\mathcal{E}_{AB}$ for the SSM with four different k_m values. The most prominent observation is that the error decays exponentially for sufficiently large k_m, showing that A and B will synchronize completely as $t \to \infty$. The figure also shows that a threshold value exists for k_m below which the error does not decay over time. With the parameters considered in Fig. 1, the threshold for SSM is $k_m \approx 5$. The symbols in Fig. 1 represent $\mathcal{E}_{AB}$ for the DSM (empty circles) and the DMM (solid circles) with $k_m = 9$. The error for DSM

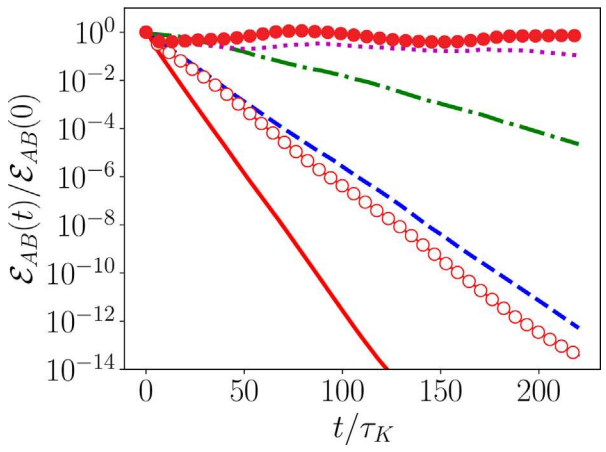

FIG. 1. The synchronization error as a function of time. Lines: SSM, Solid line: $k_m = 9$, dashed line: $k_m = 7$, dash-dotted line: $k_m = 5.5$, and dotted line: $k_m = 5$. Empty circles: DSM with $k_m = 9$ and solid circles: DMM with $k_m = 9$. For $N_{\mathrm{LES}} = 64$ and $Re_\lambda = 112$.

clearly decays much slower than that for SSM, which implies two LES, supplemented with DSM, synchronize slower than those supplemented with SSM. For $k_m = 9$, the systems with DMM as the model do not synchronize at all. Thus, the synchronization can happen with very different rates for different models.

Figure 1 shows that, for sufficiently large k_m, the synchronization error can be fitted with an exponential function

$$\mathcal{E}_{AB} \sim \exp(-\alpha t/\tau_K), \tag{15}$$

where τ_K is the Kolmogorov time scale of the DNS velocity and α is a coefficient referred to as the decay rate. α can be found as the slope of the linear regression of the data for $(t/\tau_K, -\log \mathcal{E}_{AB})$. The values of α for different k_m and filter scale Δ, or equivalently the cutoff wavenumber $k_\Delta = \pi/\Delta$, have been found and plotted in Fig. 2. We focus on k_m around the threshold value $k_c \approx 0.2\eta^{-1}$ reported in Yoshida, Yamaguchi, and Kaneda.[11] For clarity of the figures, only results for $N_{\mathrm{LES}} = 64$ and 96 are presented. The results show that, as expected, the decay rate increases with $k_m\eta$ for all models. However, other parameters being equal, the decay rate for SSM is the largest, the one for DSM is the second, and the one for DMM is the smallest. The results for different Reynolds numbers are distinguished with symbols of different shapes, whereas symbols of the same shape but with different sizes indicate different k_Δ. Therefore, it is clear that, for SSM and DSM, α increases with Re_λ but decreases with k_Δ. However, for DMM, α shows no clear dependence on k_Δ or Re_λ.

The k_m value for which $\alpha = 0$ is the threshold wavenumber k_c. If $k_m < k_c$, then synchronization between the two LES (in systems A and B, respectively) is impossible. To find k_c in a given case, we start from a k_m value for which $\mathcal{E}_{AB}$ decays exponentially over time. The value of k_m is then decremented by 0.5 repeatedly, with $\mathcal{E}_{AB}(t)$ calculated in each case, until a k_m value is found for which $\mathcal{E}_{AB}(t)$ does not decay. The average of the final two values of k_m is taken to be k_c. This means that the true value of k_c falls between $k_c \pm 0.25$.

The normalized threshold wavenumber $k_c\eta$ is plotted against Δ/η in Fig. 3(a) and $k_\Delta\eta$ in Fig. 3(b). The error bars indicate the values corresponding to $k_c \pm 0.25$. Consistent with Fig. 2, $k_c\eta$ increases with $k_\Delta\eta$ for SSM and DSM, while it undulates around a value close to 0.2 for DMM. For large $k_\Delta\eta$, $k_c\eta$ approaches, the DNS value 0.2, which is to be expected. For a given $k_\Delta\eta$, the threshold is the smallest for SSM, somewhat larger for DSM, and the largest for DMM.

The results in Figs. 2 and 3 depict a consistent picture: the rate of synchronization and the data needed to synchronize two identical LES with DMM are roughly the same as those needed to synchronize two identical DNS, whereas the synchronization of LES with SSM or DSM requires less data and can be achieved quicker. In the case with the smallest $k_\Delta\eta$ (hence, largest Δ/η) shown in Fig. 3, the ratio between k_c for SSM and DMM is approximately $0.13/0.21$, which means the data needed for SSM is only $(0.13/0.21)^3 \approx 24\%$ of those for DMM. This is a significant saving.

The slave systems are synchronized because their conditional Lyapunov exponent (with the assimilated data given) is negative.[20] As SSM and DSM are slightly more dissipative than DMM, it is perhaps not surprising that the LES with these two models has smaller conditional Lyapunov exponents, and hence is easier to synchronize.

The results in Fig. 3 suggest that the values of $k_c\eta$ for SSM and DSM may decay and asymptote toward small constants as $k_\Delta\eta$ decreases further. However, simulations with much smaller η are needed to ascertain the behavior, which are beyond our computational capacity at the moment.

The fact that less data are needed to synchronize the LES with SSM and DSM as the SGS models to DNS do not immediately suggest

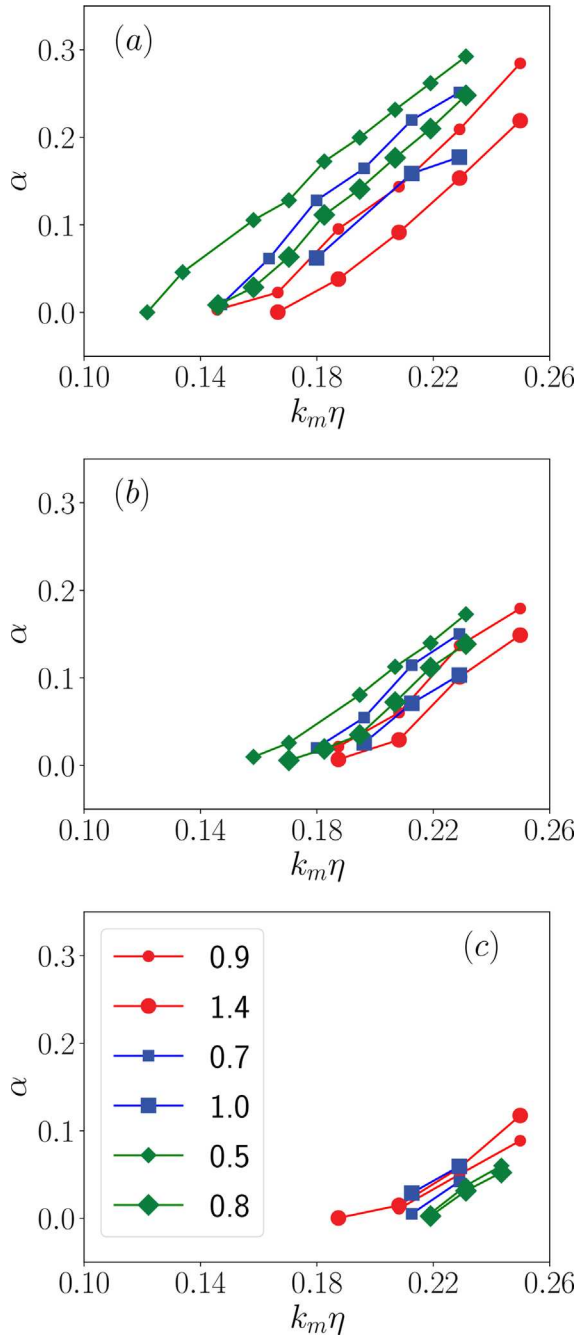

FIG. 2. The decay rate for (a) SSM, (b) DSM, and (c) DMM as a function of $k_m\eta$. Circles: $Re_\lambda = 75$, squares: $Re_\lambda = 90$, and diamonds: $Re_\lambda = 112$. Larger symbols: $N_{\mathrm{LES}} = 96$ and smaller symbols: $N_{\mathrm{LES}} = 64$. The inset in (c) shows the values of $k_\Delta\eta$ corresponding to the curves in (a)–(c).

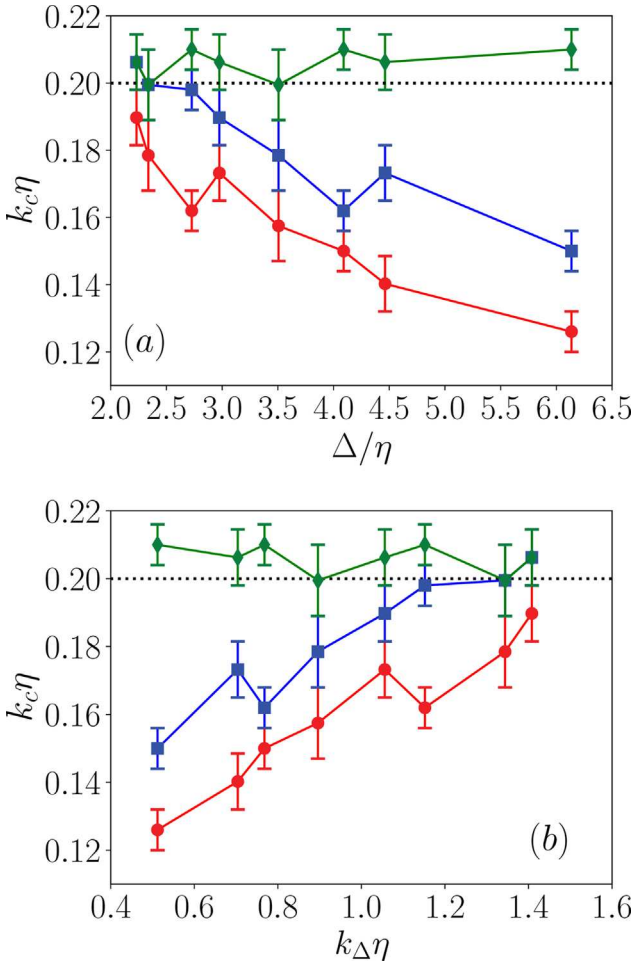

FIG. 3. The threshold wavenumber $k_c\eta$ as a function of (a) Δ/η and (b) $k_\Delta\eta$. Circles: SSM, squares: DSM, and diamonds: DMM. The error bars represent $k_c\eta \pm 0.25\eta$.

that the LES velocity obtained with SSM or DSM is more accurate compared to those obtained with DMM. This issue is addressed next.

B. The error in the synchronized LES velocity

We turn to the second objective, i.e., the relationship between the DNS velocity (the master system M) and the LES velocity (the slave systems). In this part, k_m is always chosen to be higher than the threshold value k_c, and only one LES, denoted as system S, is simulated simultaneously with the DNS (master system M). However, for each set of parameters, five simulations (solving concurrent systems M and S) are conducted, and the statistics of the errors are averaged over the five realizations.

Let $\overline{\boldsymbol{u}}$ be the LES velocity in system S. We will need to compare $\overline{\boldsymbol{u}}$ with filtered DNS data. We use $\overline{\boldsymbol{v}}$ to denote the filtered DNS velocity field. $\overline{\boldsymbol{v}}$ is obtained from the DNS velocity in the master system M by filtering the latter with a cutoff filter[24] with cutoff wavenumber k_Δ.

Our focus will be the point-wise difference between instantaneous velocities $\overline{\boldsymbol{u}}$ and $\overline{\boldsymbol{v}}$. Nevertheless, we will also document the energy spectra of the two velocity fields. Let $\hat{\overline{\boldsymbol{u}}}(\boldsymbol{k},t)$ and $\hat{\overline{\boldsymbol{v}}}(\boldsymbol{k},t)$ be the Fourier coefficients of $\overline{\boldsymbol{u}}$ and $\overline{\boldsymbol{v}}$, respectively. Let $E_v(k,t)$ and $E_u(k,t)$ denote the energy spectra of $\overline{\boldsymbol{v}}$ and $\overline{\boldsymbol{u}}$ at time t, respectively, where

$$E_v(k,t) = \frac{1}{2}\oint_{S_k} \langle|\hat{\overline{\boldsymbol{v}}}(\boldsymbol{k},t)|^2\rangle dS_k \tag{16}$$

and

$$E_u(k,t) = \frac{1}{2}\oint_{S_k} \langle|\hat{\overline{\boldsymbol{u}}}(\boldsymbol{k},t)|^2\rangle dS_k, \tag{17}$$

where $\langle\ \rangle$ indicates ensemble average over the five realizations. To characterize the point-wise difference between $\overline{\boldsymbol{u}}$ and $\overline{\boldsymbol{v}}$, a suitable statistic is the energy spectrum of $\overline{\boldsymbol{u}} - \overline{\boldsymbol{v}}$. The spectrum of $\overline{\boldsymbol{u}} - \overline{\boldsymbol{v}}$ is defined by

$$E_d(k,t) = \frac{1}{2}\oint_{S_k} \langle|\hat{\overline{\boldsymbol{u}}}(\boldsymbol{k},t) - \hat{\overline{\boldsymbol{v}}}(\boldsymbol{k},t)|^2\rangle dS_k. \tag{18}$$

Since the Fourier modes with $|\boldsymbol{k}| \le k_m$ in $\overline{\boldsymbol{u}}$ and $\overline{\boldsymbol{v}}$ are the same due to data assimilation, $E_d(k,t) = 0$ for $k \le k_m$. The total error in $\overline{\boldsymbol{u}}$ compared with $\overline{\boldsymbol{v}}$, denoted by $\mathcal{E}(t)$, can be found from E_d according to

$$\mathcal{E}(t) = \int_0^{k_\Delta} E_d(k,t)dk = \int_{k_m}^{k_\Delta} E_d(k,t)dk. \tag{19}$$

$\mathcal{E}(t)$ measures the error in the Fourier modes with $|\boldsymbol{k}| > k_m$. In other words, it measures the error in the scales that are not affected by DA directly.

We also introduce a spectral correlation function

$$\rho(k,t) = \frac{1}{2\sqrt{E_u E_v}}\oint_{S_k} \langle\hat{\overline{\boldsymbol{u}}}(-\boldsymbol{k},t)\cdot\hat{\overline{\boldsymbol{v}}}(\boldsymbol{k},t)\rangle dS_k. \tag{20}$$

Note that $\rho(k,t) = 1$ for $k \le k_m$ because $\hat{\overline{\boldsymbol{u}}}(\boldsymbol{k},t) = \hat{\overline{\boldsymbol{v}}}(\boldsymbol{k},t)$ when $|\boldsymbol{k}| \le k_m$. $\rho(k,t)$ measures the correlation between the Fourier modes of $\overline{\boldsymbol{u}}$ and $\overline{\boldsymbol{v}}$, and hence complements the information provided by $E_d(k,t)$ and $\mathcal{E}(t)$. With $\rho(k,t)$, a *correlation wavenumber* can be defined

$$k_\rho(t) = \int_0^{k_\Delta} \rho(k,t)dk. \tag{21}$$

$k_\rho(t)$ gives a measure of the range of wavenumbers in which the two fields are well correlated. Using the fact that $\rho(k,t) = 1$ for $k \le k_m$, we have

$$k_\rho(t) = k_m + \int_{k_m}^{k_\Delta} \rho(k,t)dk. \tag{22}$$

Thus, we can introduce a dimensionless correlation wavenumber

$$k_\rho^*(t) = \frac{k_\rho - k_m}{k_\Delta - k_m} = \frac{1}{k_\Delta - k_m}\int_{k_m}^{k_\Delta} \rho(k,t)dk, \tag{23}$$

which presents $k_\rho(t)$ as a fraction of the full range of wavenumbers above k_m. Obviously, $0 \le k_\rho^* \le 1$ and a larger k_ρ^* indicates better agreement between $\overline{\boldsymbol{u}}$ and $\overline{\boldsymbol{v}}$.

The parameters introduced above are time dependent. We will use subscript T to denote their time-averaged values. Thus, $E_{uT}(k)$ is the time-averaged energy spectrum for $\overline{\boldsymbol{u}}$. Similarly, we have

parameters $E_{vT}(k)$, $E_{dT}(k)$, $\mathcal{E}_T$, and $k^*_{\rho T}$ which are the time-averaged values of E_v, E_d, $\mathcal{E}$, and k^*_ρ, respectively.

Finally, we let

$$K_\Delta \equiv \int_{k_m}^{k_\Delta} E_v(k, t=0)dk, \tag{24}$$

where $t=0$ is the time when DA is initiated. K_Δ is used to normalize $\mathcal{E}(t)$ and $\mathcal{E}_T$.

We are ready to present the analyses of the data. We start with the contours of the velocity fields shown in Fig. 4 which provide an intuitive visualization of the effects of data assimilation. The two distributions display a remarkable similarity in their large-scale structures, as a consequence of data assimilation. The LES field, shown in Fig. 4(a), misses some of the small-scale features that we can observe in the DNS field, shown in Fig. 4(b). This is to be expected as LES does not resolve all the scales in the DNS fields. Our main focus is to quantify the errors in the small scales of the LES fields.

We now present selected results for $\mathcal{E}(t)$ to illustrate some general quantitative features. Figure 5 plots $\mathcal{E}(t)/K_\Delta$ for $Re_\lambda = 112$ and $k_m\eta = 0.21$. The figure shows that $\mathcal{E}(t)$ decreases over time and then levels off. Therefore, the data assimilated at large scales propagate toward the small scales and push the small scales of the LES velocity fields toward the DNS velocity. The stationary value of $\mathcal{E}(t)$ decreases with increasing $k_\Delta\eta$ for all three models. Interestingly, the errors for SSM and DSM are significantly smaller than that for DMM for the case with $k_\Delta\eta = 0.5$. DSM and DMM only become more accurate than SSM when $k_\Delta\eta$ is increased.

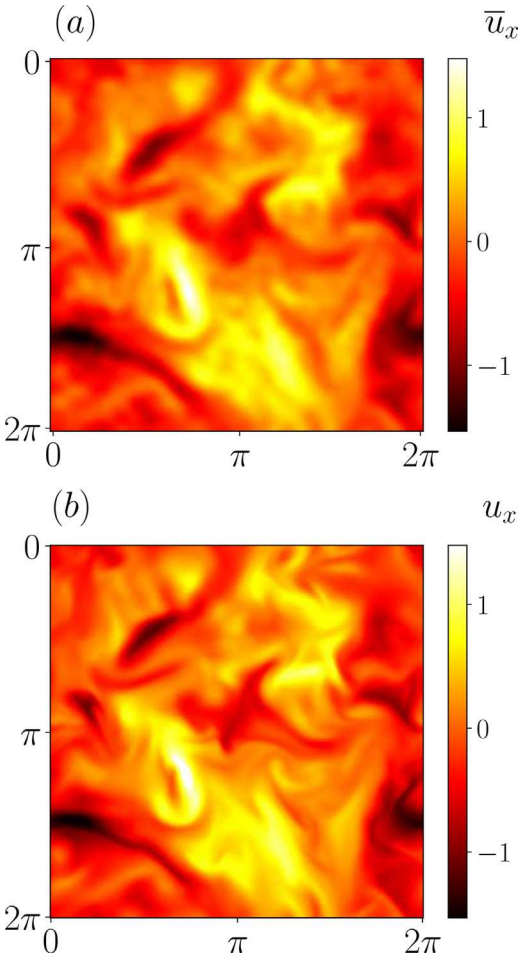

FIG. 4. The instantaneous distribution of (a): $\overline{u}_x$ from LES with SSM and (b): v_x from DNS. $Re_\lambda = 112$, $k_\Delta\eta = 0.5$, $k_m\eta = 0.21$.

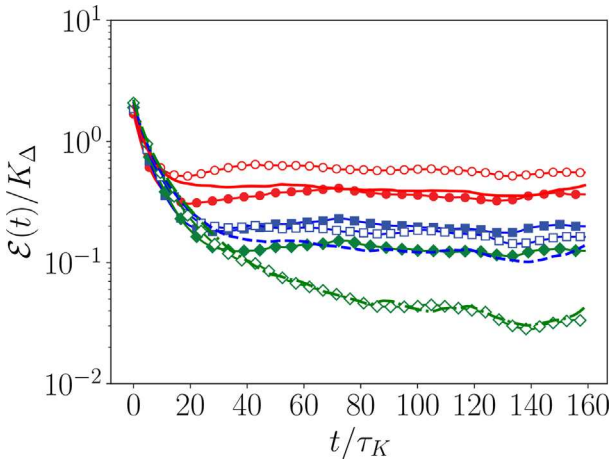

FIG. 5. The normalized error $\mathcal{E}(t)/K_\Delta$. Solid symbols: SSM, lines: DSM, and empty symbols: DMM. $k_\Delta\eta = 0.5$: circles and solid line; $k_\Delta\eta = 0.8$: squares and dashed line; $k_\Delta\eta = 1.0$: diamonds and dotted line. For $k_m\eta = 0.21$ and $Re_\lambda = 112$ only.

The data assimilation process does affect the small scales of the instantaneous LES velocity fields. However, it does not visually alter their *statistics*. This observation is demonstrated in Fig. 6 with the time-averaged energy spectra. The spectra are obtained for the cases where $k_m\eta = 0.21$ which corresponds to $k_m = 9$. As a result, the LES spectra agree with the DNS one for $k_m < 9$ simply due to the fact that the Fourier modes in the LES fields are exact copies of those in the

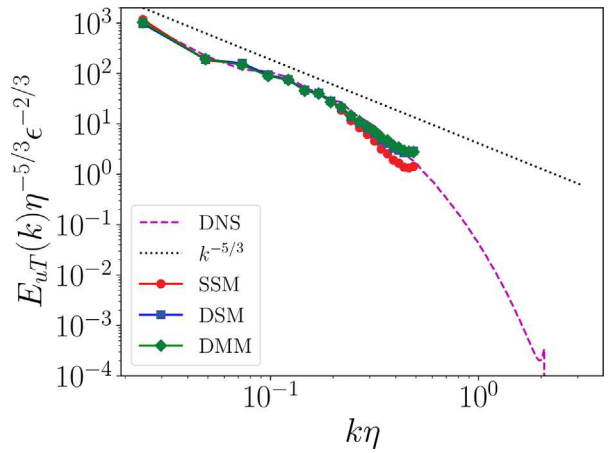

FIG. 6. Energy spectra for SSM (circles), DSM (squares), and DMM (diamonds) compared with the DNS energy spectrum (dashed line), for $Re_\lambda = 112$, $k_m\eta = 0.21$, and $N_{LES} = 64$ only. Dotted line: the $-5/3$ Kolmogorov spectrum.

DNS field. The LES energy spectra for k approaching k_Δ display well-known features. Namely, we observe the SSM underestimates the high wavenumber end of the spectrum. On the other hand, DSM and DMM reproduce the DNS spectrum quite well, due to their ability to adjust the model coefficients via the dynamic procedure. This result illustrates the point we mentioned above that DA does not significantly alter the statistics produced by the SGS models. However, as we have demonstrated briefly in Fig. 5 and will investigate in more detail later, it can significantly improve the prediction of the instantaneous point-wise distribution of the velocity fields, which is the main benefit of DA in this case.

We now look into the statistics of the instantaneous difference between $\overline{\boldsymbol{u}}$ and $\overline{v}$ in more detail. The time-averaged error $\mathcal{E}_T$ and the time-averaged dimensionless correlation wavenumber $k^*_{\rho T}$ are plotted in Figs. 7 and 8. Figure 7(a) shows, first, for a given $k_m\eta$, the correlation between the LES and the DNS fields increases with $k_\Delta\eta$. This behavior is to be expected as the LES is more accurate for larger $k_\Delta\eta$. Second, for a given $k_\Delta\eta$, $k^*_{\rho T}$ does not always increase with $k_m\eta$.

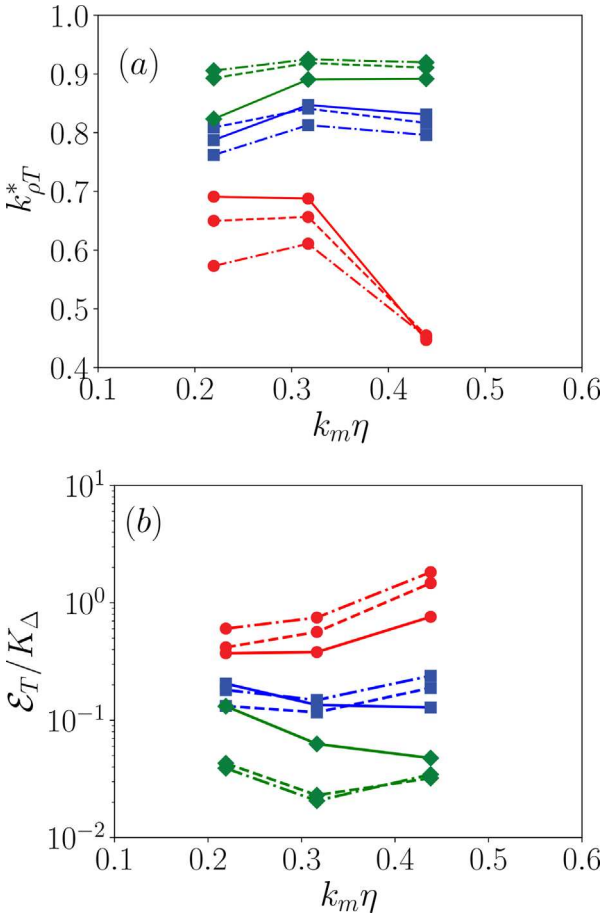

FIG. 7. (a): Dimensionless correlation wavenumber $k^*_{\rho T}$ as a function of $k_m\eta$. (b): $\mathcal{E}_T/K_\Delta$ as a function of $k_m\eta$. For $Re_\lambda = 112$. Circles: $k_\Delta\eta = 0.5$, squares: $k_\Delta\eta = 0.8$, and diamonds: $k_\Delta\eta = 1.0$. Solid lines: SSM, dash lines: DSM, and dash-dotted lines: DMM.

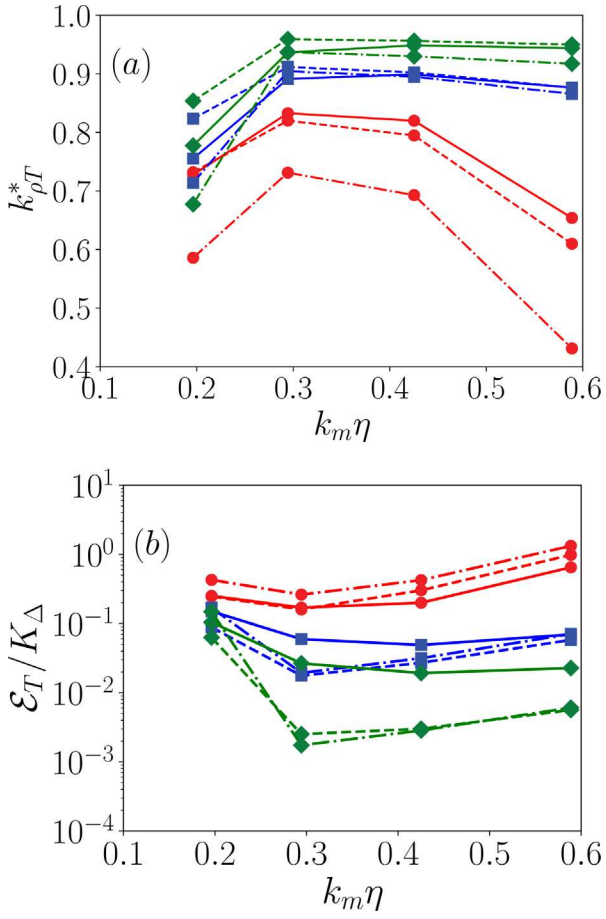

FIG. 8. Same as Fig. 7 but for $Re_\lambda = 90$. (a): Dimensionless correlation wavenumber $k^*_{\rho T}$ as a function of $k_m\eta$. (b): $\mathcal{E}_T/K_\Delta$ as a function of $k_m\eta$.

This occurs because $k^*_{\rho T}$ measures only the correlation of the Fourier modes with $k_m \le |\boldsymbol{k}| \le k_\Delta$. For large k_m, the correlation is predominantly determined by motions with wavenumbers close to k_Δ which have larger errors due to the errors in the SGS models, thus degrading the correlation. This effect is not obvious when k_Δ is sufficiently large compared with k_m, but it is rather pronounced for the cases where $k_\Delta\eta = 0.5$. Third, for the case with the smallest $k_\Delta\eta$, $k^*_{\rho T}$ for SSM is the highest among the three models (at least if $k_m\eta$ is not too large). Thus, for larger filter scales, LES with SSM synchronizes with DNS better than LES with other models.

The above observations are consistent with the result for the mean error $\mathcal{E}_T/K_\Delta$ shown in Fig. 7(b). Notably, the normalized error for SSM is the smallest for $k_\Delta\eta = 0.5$, which is about 30% at $k_m\eta = 0.21$ as opposed to approximately 60% for DMM. DMM becomes more accurate for larger $k_\Delta\eta$. Arguably, DSM displays the best overall performance. Figure 8 shows the results for $Re_\lambda = 90$, which depicts similar trends.

The spectral distribution of the error in the LES velocity is shown in Fig. 9, where $E_{dT}(k)/E_{vT}(k)$ is plotted against a normalized wavenumber

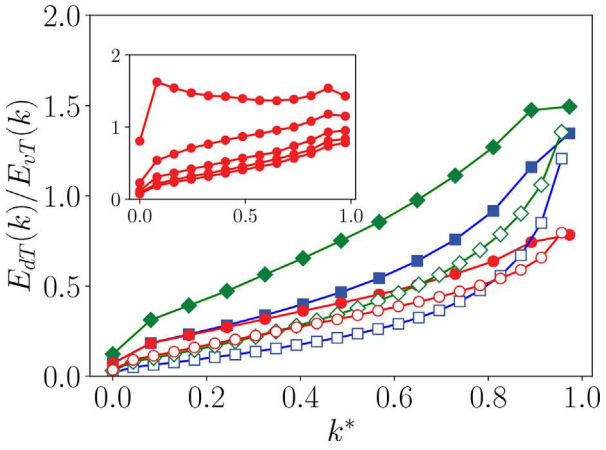

FIG. 9. The normalized time-averaged spectral error $E_{dT}(k)/E_{vT}(k)$ as a function of the normalized wavenumber k^*. Circles: SSM, squares: DSM, and diamonds: DMM. Solid symbols: $k_\Delta\eta = 0.5$ and empty symbols: $k_\Delta\eta = 0.8$. Inset: the spectral error $E_u(k,t)/E_{vT}(k)$ for the SSM at different times where, from top to bottom, $t/\tau_K = 0, 3.62, 7.25, 10.87, 14.50$. For $k_m\eta = 0.21$ and $Re_\lambda = 112$.

$$k^* = \frac{k - k_m}{k_\Delta - k_m}.$$

For clarity, only results for $k_\Delta\eta = 0.5$ and 0.8 are plotted. Interestingly, for $k_\Delta\eta = 0.5$, the synchronization error in the velocity obtained with SSM is the smallest across all wavenumbers, which explains the observation from Fig. 7 that the total error is the smallest. For $k_\Delta\eta = 0.8$, Fig. 9 shows that the error for the DSM is the smallest for k^* values up to 0.8, but it becomes bigger than the one for SSM for $k^* > 0.8$. The results for $k_\Delta\eta = 1.0$ are not shown, but they also display similar features.

The inset plots $E_d(k,t)/E_{vT}(k)$ for a selected case, which illustrates how the error spectrum changes over time. As is expected, the error at a given k decreases over time, which shows that the effects of data assimilation propagate from large scales to small scales.

C. The synchronization of the vorticity field

The vorticity field provides insights into the small scales of the velocity field. It is instructive to examine the synchronization of the vorticity since it might be affected more significantly by the synchronization error in the small scales.

We let $\overline{\boldsymbol{\omega}}^s \equiv \nabla \times \overline{\boldsymbol{u}}$ be the vorticity of the LES velocity and use $\overline{\boldsymbol{\omega}}^m \equiv \nabla \times \overline{\boldsymbol{v}}$ to denote the filtered vorticity calculated from filtered DNS data. The strength of the vorticity fields is represented by the normalized entrophy

$$\Omega^s = \frac{\overline{\omega}_i^s \overline{\omega}_i^s}{\langle \overline{\omega}_i^m \overline{\omega}_i^m \rangle}, \quad \Omega^m = \frac{\overline{\omega}_i^m \overline{\omega}_i^m}{\langle \overline{\omega}_i^m \overline{\omega}_i^m \rangle}, \tag{25}$$

for the LES and the filtered DNS vorticity, respectively. Note that both are normalized by the same value $\langle \overline{\omega}_i^m \overline{\omega}_i^m \rangle$.

The instantaneous distributions of Ω^s and Ω^m are illustrated in Figs. 10–12 with 2D slices of the fields at arbitrary chosen times in the stationary stage. The figures depict the prevalence of concentrated strong vortical structures. It is observed that LES, regardless of the SGS model being used, reproduces the instantaneous shapes and locations

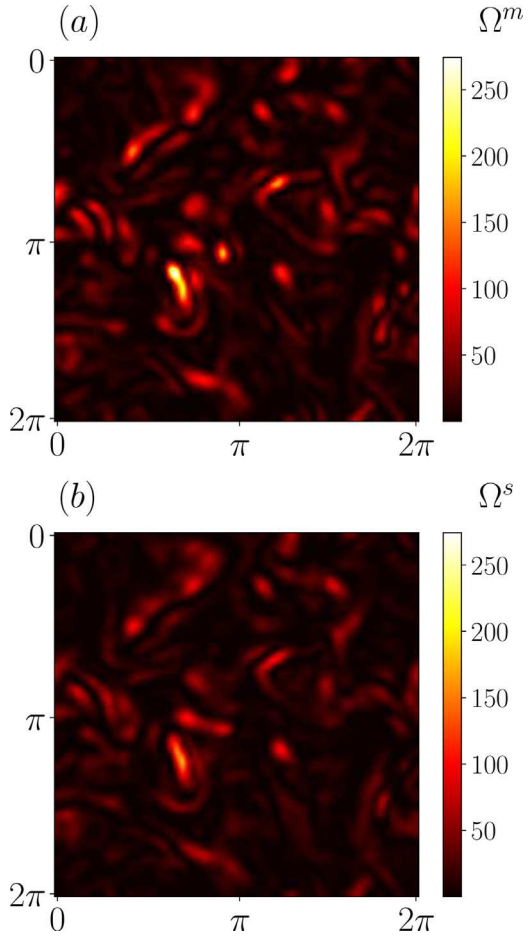

FIG. 10. The instantaneous distribution of the normalized entrophy from (a): filtered DNS data and (b): LES data with SSM. $Re_\lambda = 112$, $k_\Delta\eta = 0.5$, $k_m\eta = 0.21$.

of the structures very well. The SSM appears to underestimate the strength of the vortical structures and fewer strong structures are observed. On the other hand, DSM and DMM appear to somewhat overestimate the entrophy.

The above qualitative observations are confirmed by the joint probability density functions (PDFs) between Ω^s and Ω^m, which are shown in Fig. 13, and the correlation coefficients ρ_ω in Table II, where ρ_ω is defined by

$$\rho_\omega = \frac{\langle (\Omega^s - \langle \Omega^s \rangle)(\Omega^m - \langle \Omega^m \rangle) \rangle}{\sqrt{(\Omega^s - \langle \Omega^s \rangle)^2 (\Omega^m - \langle \Omega^m \rangle)^2}}. \tag{26}$$

The correlation is very high in all cases, and they are stronger, for all three models, when $k_\Delta\eta$ is larger. For LES with the coarser grids ($k_\Delta\eta = 0.5$), the SMM result has the highest correlation with the filtered DNS data, whereas the best correlation is observed for DSM for $k_\Delta\eta = 0.8$. The correlation coefficients capture the correlation between fluctuations of Ω^s and Ω^m at the order of the magnitude of their standard deviations. Correlations between large fluctuations are illustrated by their joint PDFs. Figure 13(a) shows the joint PDFs

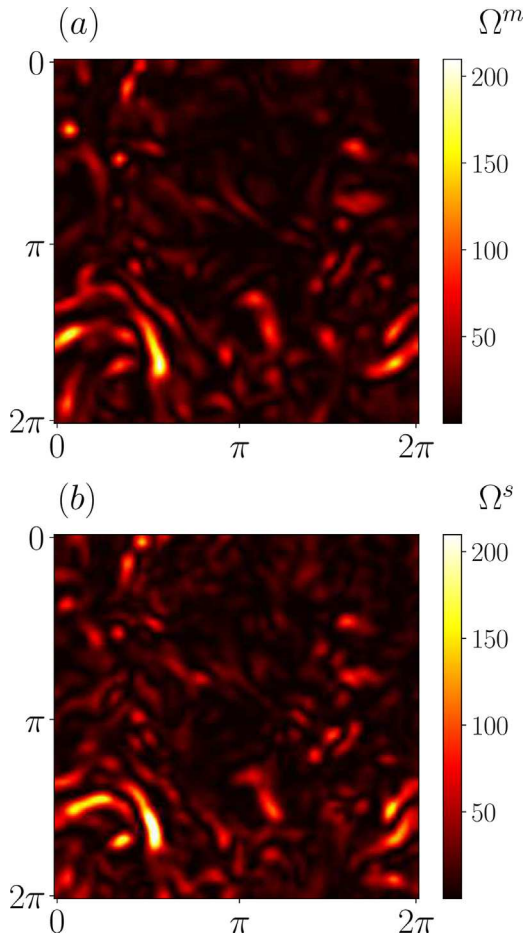

FIG. 11. The instantaneous distribution of the normalized entropy from (a): filtered DNS data and (b): LES data with DSM. $Re_\lambda = 112$, $k_\Delta\eta = 0.5$, $k_m\eta = 0.21$.

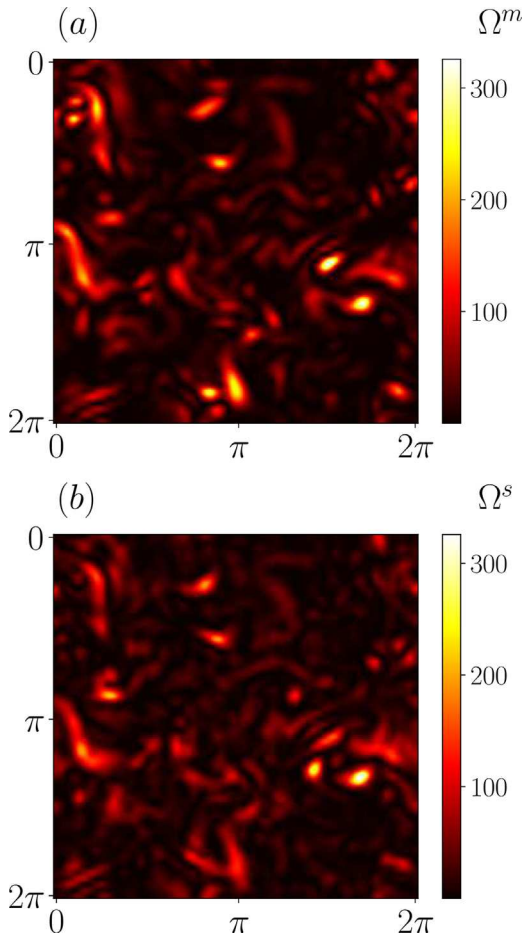

FIG. 12. The instantaneous distribution of the normalized entropy from (a): filtered DNS data and (b): LES data with DMM. $Re_\lambda = 112$, $k_\Delta\eta = 0.5$, $k_m\eta = 0.21$.

calculated for the three models at $k_\Delta\eta = 0.5$. Compared with the contours of the joint PDFs calculated with DSM and DMM, those calculated with SSM are more elongated along the diagonal, which suggests that the correlations for larger fluctuations are also stronger for the SSM. The contours for SSM tilt toward the lower side of the diagonal, indicating that SSM tends to underestimate the large fluctuations, which is consistent with Fig. 10. The contours for DSM and DMM are mostly symmetrical with respect to the diagonal, which shows that there is systematic under- or overestimation. However, the contours spread wider around the diagonal, which explains why the correlations for these two models are smaller than that for SSM. For $k_\Delta\eta = 0.8$, the joint PDFs are shown in Fig. 13(b). The joint PDF for DSM clearly indicates the strongest correlation between Ω^s and Ω^m since the contours align with the diagonal much better than those of the joint PDFS for SSM and DMM.

The alignment between the vorticity vectors $\overline{\boldsymbol{\omega}}^m$ and $\overline{\boldsymbol{\omega}}^s$ is another indicator of the synchronization between the LES and the DNS velocity fields. Figure 14 plots the PDF of $\cos(\overline{\boldsymbol{\omega}}^m, \overline{\boldsymbol{\omega}}^s)$, where $(\overline{\boldsymbol{\omega}}^m, \overline{\boldsymbol{\omega}}^s)$ denotes the angle between $\overline{\boldsymbol{\omega}}^m$ and $\overline{\boldsymbol{\omega}}^s$. Perfect alignment corresponds to $\cos(\overline{\boldsymbol{\omega}}^m, \overline{\boldsymbol{\omega}}^s) = 1$. Figure 14 shows that the LES vorticity is strongly aligned with the filtered DNS vorticity, which shows that assimilation of large-scale velocity does improve the prediction of vorticity fields. For $k_\Delta\eta = 0.5$, the alignment is the strongest for SSM, consistent with previous results. For $k_\Delta\eta = 0.8$, the alignment improves for all models. The alignment obtained with DSM becomes the strongest, while the alignment obtained with DMM is still the worst.

Vortex stretching is the main mechanism that controls the strength of the vorticity field. Mathematically the effect is captured by the so-called vortex stretching term. For LES, this term is given by

$$P^s_\omega \equiv \overline{\omega}^s_i \overline{s}_{ij} \overline{\omega}^s_j = \overline{\omega}^s_i \overline{\omega}^s_j \partial_j \overline{u}_i, \tag{27}$$

which can be calculated from the LES velocity $\overline{\boldsymbol{u}}$. For DNS, the vortex stretching term can be calculated from the same expression, with $\overline{\boldsymbol{u}}$ replaced the filtered velocity $\overline{\boldsymbol{v}}$ and $\overline{\boldsymbol{\omega}}^s$ replaced by $\overline{\boldsymbol{\omega}}^m$. We calculate the correlation coefficient ρ_P of P^s_ω and P^m_ω, where

$$\rho_P = \frac{\langle (P^s_\omega - \langle P^s_\omega \rangle)(P^m_\omega - \langle P^m_\omega \rangle) \rangle}{\sqrt{(P^s_\omega - \langle P^s_\omega \rangle)^2 (P^m_\omega - \langle P^m_\omega \rangle)^2}} \tag{28}$$

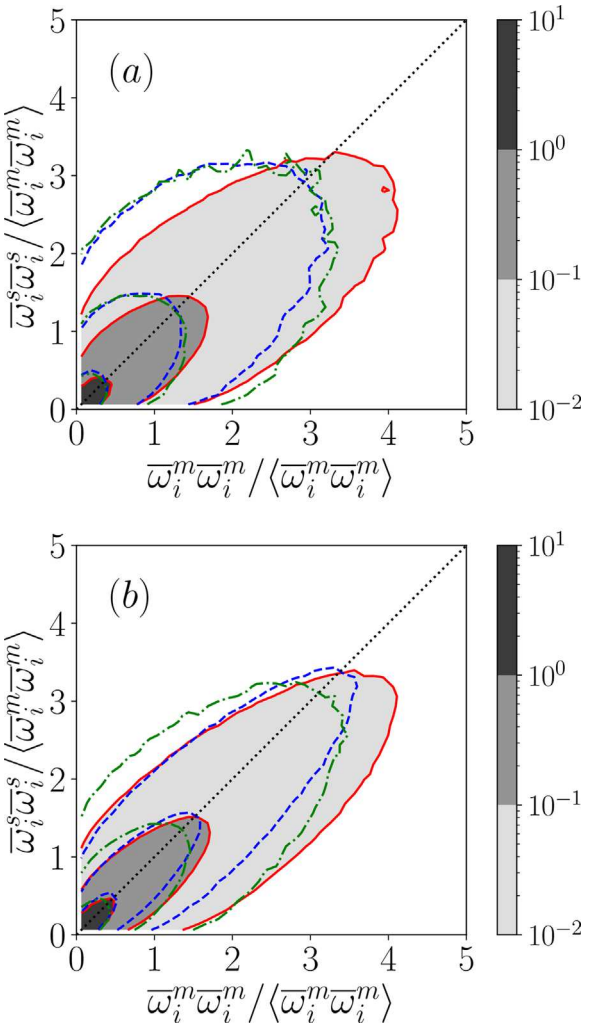

FIG. 13. The joint PDFs of the normalized entrophy for (a): $k_\Delta\eta = 0.5$ and (b): $k_\Delta\eta = 0.8$. Gray scale and red lines: SSM, dashed lines: DSM, and dash-dotted lines: DMM. Values of the contours: 10^{-2}, 10^{-1}, 1 (started from outermost) $Re_\lambda = 112$, $k_m\eta = 0.21$.

and their joint PDFs. The results are shown in Fig. 15 for $k_\Delta\eta = 0.5$ and 0.8.

For the smaller k_Δ (hence, larger Δ), Fig. 15(a) shows that the correlation is the strongest for SSM with $\rho_P = 0.75$, and slightly weaker for DSM with $\rho_P = 0.64$. With $\rho_P = 0.58$, the correlation for DMM is the weakest. The contours of the joint PDF for SSM are tilted toward the horizontal axis, which shows that SSM somewhat

TABLE II. Correlation coefficient ρ_ω for the normalized entrophy.

$k_\Delta\eta$	SSM	DSM	DMM
0.5	0.87	0.81	0.78
0.8	0.89	0.93	0.86

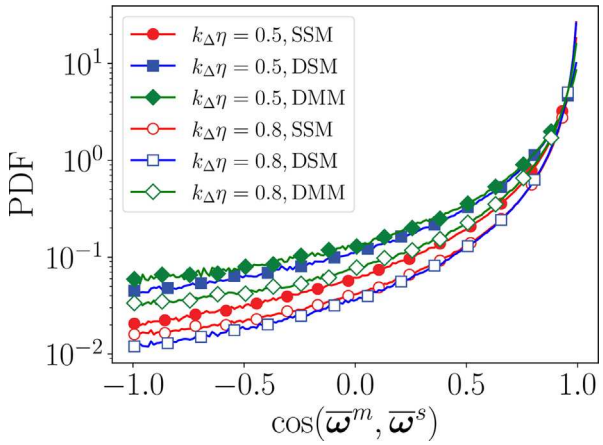

FIG. 14. The PDF of $\cos(\overline{\omega}^m, \overline{\omega}^s)$ with $Re_\lambda = 112$ and $k_m\eta = 0.21$. Solid symbols: $k_\Delta\eta = 0.5$ and empty symbols: $k_\Delta\eta = 0.8$. Circles: SSM, squares: DSM, and diamonds: DMM.

underpredicts the large fluctuations in vortex stretching. On the other hand, DSM and DMM tend to overpredict it. Overall, the joint PDF for SSM concentrates around the diagonal better than the joint PDF for the other two models, which is consistent with the observed higher correlation for SSM.

The correlation between LES predictions and DNS values for the vortex stretching term is increased when Δ is decreased, as shown in Fig. 15(b) by the values of ρ_P as well as the shapes of the joint PDFs. The vortex stretching term calculated with DSM now displays the strongest correlation with its counterpart in DNS. The result calculated with SSM is still better correlated with the DNS data than the result calculated with DMM, which is consistent with previous results (c.f. Fig. 14). The joint PDFs show that SSM still tends to underestimate the large fluctuations in P^s_ω, whereas DMM tends to overestimate them. Meanwhile, this systematic deviation is not observed in the joint PDF for DSM, as the contours are mostly symmetrical with respect to the diagonal.

D. Effects of noise

We have so far assumed the data being assimilated into the LES velocity are obtained from DNS, hence, noiseless. In practical applications of data assimilation, the data very often are obtained from experiments, and hence contain errors. To understand how the synchronization between the LES and DNS velocities might be changed in those circumstances, we conduct a set of numerical experiments with noise introduced into the data before they are assimilated.

The noise is introduced through a phase-scrambling process. Recall that $\hat{\overline{\boldsymbol{v}}}(\boldsymbol{k}, t)$ is the Fourier coefficient of the DNS velocity with wavenumber $\boldsymbol{k}$ and $\hat{\overline{\boldsymbol{u}}}(\boldsymbol{k}, t)$ is that of the LES velocity. We may write

$$\hat{\overline{\boldsymbol{v}}}(\boldsymbol{k}, t) = |\hat{\overline{\boldsymbol{v}}}(\boldsymbol{k}, t)| e^{i\phi(\boldsymbol{k}, t)}, \tag{29}$$

where $\phi(\boldsymbol{k}, t)$ is the phase angle of $\hat{\overline{\boldsymbol{v}}}$ (ϕ is also known as the argument). To introduce noise into the assimilated data, instead of letting $\hat{\overline{\boldsymbol{u}}}(\boldsymbol{k}, t) = \hat{\overline{\boldsymbol{v}}}(\boldsymbol{k}, t)$, we let

$$\hat{\overline{\boldsymbol{u}}}(\boldsymbol{k}, t) = |\hat{\overline{\boldsymbol{v}}}(\boldsymbol{k}, t)| e^{i\phi(\boldsymbol{k}, t)(1+d_\phi)} \tag{30}$$

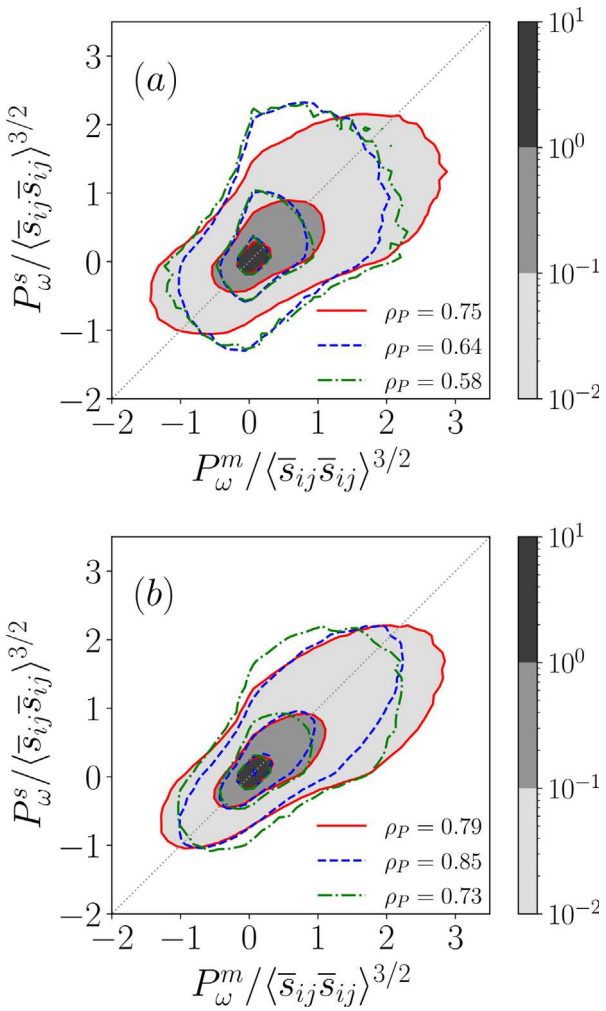

FIG. 15. The joint PDF of the vortex stretching terms with (a) $k_\Delta\eta = 0.5$ and (b) $k_\Delta\eta = 0.8$. Solid lines and gray scales: SSM, dashed lines: DSM, and dash-dotted lines: DMM. Values of the contours: 10^{-2}, 10^{-1}, 1 (started from outermost). Inset: correlation coefficients. $Re_\lambda = 112$ and $k_m\eta = 0.21$.

at each time step for $|\boldsymbol{k}| \le k_m$, where d_ϕ is a uniformly distributed random number in the interval $[-D_\phi, D_\phi]$ for some given D_ϕ. d_ϕ can be different for different $\boldsymbol{k}$ and t. Equation (30) shows that the phase angles of different Fourier modes are perturbed randomly before the Fourier modes are assimilated into the LES velocity field.

Apart from assimilating the data using Eq. (30), the rest of the numerical experiments is essentially the same as before. We consider only DNS with 256^3 grid points and $k_m\eta = 0.21$, but LES with two resolutions $N_{\rm LES} = 64$ and $N_{\rm LES} = 96$ are tested. $D_\phi = 0.2$ is used, which means that the phase angle ϕ can be modified by up to $\pm 20\%$ randomly. Due to the noise, the total error in the assimilated Fourier modes with $|\boldsymbol{k}| \le k_m$ is about 5% of the total kinetic energy contained in these modes. Therefore, the amplitude of the noise is rather significant.

We present two key results, those of $k^*_{\rho T}$ and $\mathcal{E}_T/K_\Delta$. The values for $k^*_{\rho T}$ are shown in Table III. The values are labeled with

TABLE III. The time-averaged dimensionless correlation wavenumber $k^*_{\rho T}$ obtained with noiseless or noisy data. $Re_\lambda = 112$. $k_m\eta = 0.21$.

	$k_\Delta\eta$	SSM	DSM	DMM
Noiseless	0.5	0.69	0.65	0.57
	0.8	0.79	0.81	0.76
Noisy	0.5	0.64	0.61	0.51
	0.8	0.70	0.71	0.68

TABLE IV. The time-averaged synchronization error $\mathcal{E}_T/K_\Delta$ obtained with noiseless or noisy data. $Re_\lambda = 112$. $k_m\eta = 0.21$.

	$k_\Delta\eta$	SSM	DSM	DMM
Noiseless	0.5	0.37	0.42	0.60
	0.8	0.20	0.13	0.18
Noisy	0.5	0.35	0.41	0.60
	0.8	0.25	0.22	0.25

"Noisy." The table also includes the values from the noiseless case as a comparison. The values for the noiseless case are the same as those in Fig. 7 (with $k_m\eta = 0.21$). The values for $\mathcal{E}_T/K_\Delta$ are collected in Table IV.

The values in the two tables show that the correlation between the LES and the DNS is reduced by the noise, but strong correlation is still observed. The correlation is degraded somewhat more strongly for larger k_Δ. For $k_\Delta\eta = 0.5$, the values of $\mathcal{E}_T/K_\Delta$ essentially have no difference with or without noise. They are increased slightly by the noise for $k_\Delta\eta = 0.8$ but remain small.

The results in these tables show that good synchronization between the LES and the DNS velocities is still achieved despite that the assimilated data contain significant errors.

E. Dependence on the forcing scheme

A natural question one may ask is how we may generalize the results to other flows. Although it is beyond the scope of this investigation to establish comprehensively the universality of the results we have obtained, we present in what follows some key results obtained with the same setup but a different forcing term. These results provide an initial assessment of the universality of the results we have obtained so far.

For this purpose, we use the linear forcing in Rosales and Meneveau,[25] where

$$\boldsymbol{f} = A_p\boldsymbol{u} \tag{31}$$

with A_p being a constant. As is commented by Rosales and Meneveau,[25] this forcing scheme is useful in the modeling of non-homogeneous problems. Besides, the forcing applies to all scales of the velocity field and, as such, is qualitatively different from the forcing in Kolmogorov flows investigated in Secs. III A–III D. In particular, the forcing term can be viewed as an elementary model for the effect of

the mean velocity gradient.[25] Therefore, a flow with this forcing scheme captures some generic features of more complex flows while, at the same time, is simple enough so that the physical mechanisms behind any new observations can be understood.

We will conduct DNS with 256^3 grid points only in this investigation, and the assimilated data are noiseless. The synchronization between the DNS fields and LES with 64^3, $96,^3$ and 128^3 grid points is considered. $A_p = 0.2$ is used, which is in the range of values tested in Rosales and Meneveau.[25] The statistics are calculated the same way as described in Sec. III B. In order to compare the results with those in Sec. III B, the parameters in Table I are used to normalize the results where appropriate.

The basic features of the simulations are illustrated by the energy spectra given in Fig. 16, which can be compared with the spectra in Fig. 6. The main observation is that the broad features of the DNS spectrum are captured by the SGS models. However, comparing Fig. 16 with Fig. 6, one can see that high wavenumber ends of the spectra calculated with the DSM and the DMM are significantly lower, showing that the models become overly dissipative in the current case. As a result, the improvement they can produce over the SSM is less significant.

The synchronization error is characterized by $k^*_{\rho T}$ and $\mathcal{E}_T$ (c.f. Fig. 7). $k^*_{\rho T}$ is plotted in Fig. 17. Comparing the results in Fig. 17 with those in Fig. 7, the similarity is that, in both cases, $k^*_{\rho T}$ increases with $k_m\eta$. Apart from the smallest $k_m\eta$, $k^*_{\rho T}$ also increases with k_Δ. These observations demonstrate the general effects of data assimilation persist in this new case. Meanwhile, there are some differences as well. First, now the SSM consistently produces the highest correlation, even for simulations with small $k_\Delta\eta$. The values of $k^*_{\rho T}$ for DSM and DMM at higher k_Δ are lower than those shown in Fig. 7. This observation may be explained in terms of Fig. 16 since it shows that the energy spectra produced by the DSM and DMM display a higher discrepancy with the DNS spectrum. Second, $k^*_{\rho T}$ is generally smaller than its value shown in Fig. 7. For $k_m\eta = 0.21$, the decrease in $k^*_{\rho T}$ is especially significant. This result suggests that, in this case, the threshold value of k_c might be slightly higher so that the phase synchronization between the DNS and the LES requires higher k_m to achieve (see also Fig. 18).

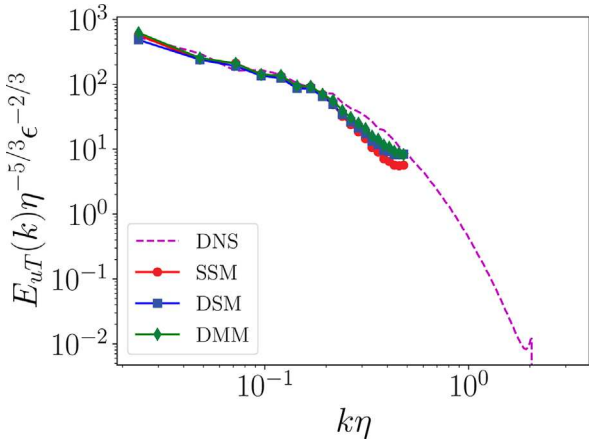

FIG. 16. The time-averaged energy spectra $E_{uT}(k)$ for $k_m\eta = 0.21$ and $N_{\mathrm{LES}} = 64$.

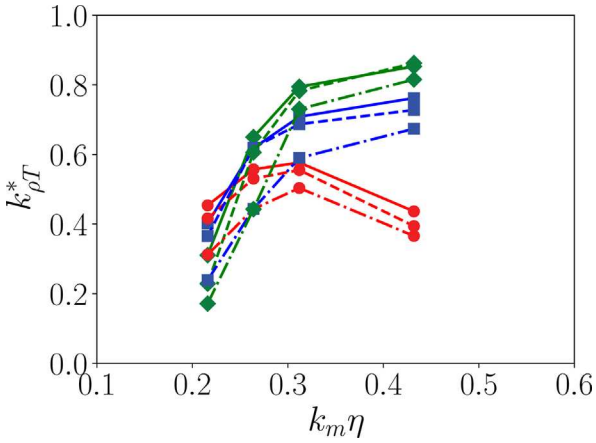

FIG. 17. The dimensionless correlation wavenumber as a function of $k_m\eta$ for simulations with linear forcing. For $Re_\lambda = 112$. Circles: $k_\Delta\eta = 0.5$, squares: $k_\Delta\eta = 0.8$, and diamonds: $k_\Delta\eta = 1.0$. Solid lines: SSM, dash lines: DSM, and dash-dotted lines: DMM.

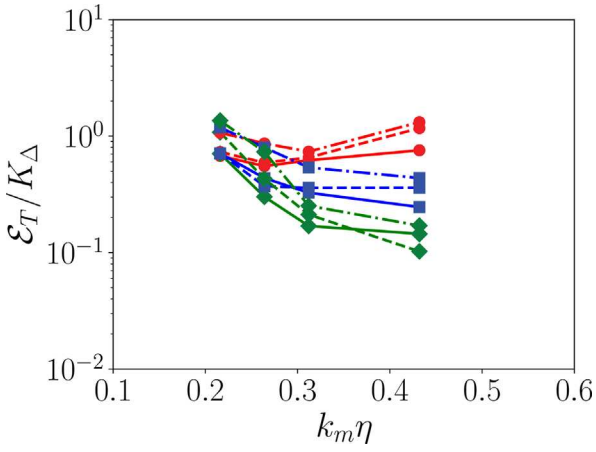

FIG. 18. The time-averaged synchronization error $\mathcal{E}_T/K_\Delta$ as a function of $k_m\eta$ for simulations with linear forcing. For $Re_\lambda = 112$. Circles: $k_\Delta\eta = 0.5$, squares: $k_\Delta\eta = 0.8$, and diamonds: $k_\Delta\eta = 1.0$. Solid lines: SSM, dash lines: DSM, and dash-dotted lines: DMM.

IV. CONCLUSIONS

We investigate the synchronization between LES and DNS velocity fields in a periodic domain when a simple data assimilation scheme is applied, where the low wavenumber modes in the former are replaced by those in the latter at each time step. The problem arises from the practice of data assimilation where measurements or experimental data are assimilated in some ways into numerical simulation dynamically to improve simulation results. LES with the standard Smagorinsky model, the dynamic Smagorinsky model, and the dynamic mixed model are conducted. Generalizations to noisy data and flows driven by different forcing terms are also investigated. The main conclusions of the analyses are as follows:

1. There is a threshold wavenumber such that LES is synchronized in phase with DNS as long as all Fourier modes with smaller

wavenumbers are assimilated, but no synchronization can be achieved if fewer Fourier modes are assimilated.
2. The threshold wavenumber is the smallest for the standard Smagorinsky model and the largest for the dynamic mixed model, whereas the one for the dynamic Smagorinsky model lies in between.
3. The threshold wavenumber for the standard and dynamic Smagorinsky models decreases as the filter scale increases, whereas the one for the dynamic mixed model does not depend on the filter scale.
4. With larger filter scales, the total synchronization error for the standard Smagorinsky model is the smallest among the models. The vorticity field also correlates with the DNS vorticity field better than other models even though strong vorticity fluctuations are somewhat underestimated. For smaller filter scales, the dynamic Smagorinsky model becomes more accurate.
5. Though there are quantitative differences, good synchronization is still obtained with noisy data or in flows with a linear forcing term. The advantages of the Smagorinsky models are still observed in these cases.

The observations suggest that, in data assimilation applications, the standard and dynamic Smagorinsky models might have advantages that have not been appreciated before. We conjecture that the better synchronization is due to the dissipative nature of these models. We should emphasize, however, that the results are obtained in a flow with periodic boundary conditions. The natural next step is to investigate the performance of the models in the simulations of wall-bounded turbulent flows with data assimilation. Wall-bounded flows present additional questions as to the roles of near wall turbulent coherent structures and the wall boundary conditions for the models. These questions will be the focus of future research.

ACKNOWLEDGMENTS

The first author acknowledges the support of the National Natural Science Foundation of China (No. 12102391) and the China Scholarship Council (No. 201808330624).

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1]E. Kalnay, *Atmospheric Modelling, Data Assimilation and Predictability* (Cambridge University Press, Cambridge, 2003).
[2]B. Protas, B. R. Noack, and J. Osth, "Optimal nonlinear eddy viscosity in Galerkin models of turbulent flows," J. Fluid Mech. **766**, 337–367 (2015).
[3]C. He, Y. Liu, and L. Gan, "A data assimilation model for turbulent flows using continuous adjoint formulation," Phys. Fluids **30**, 105108 (2018).
[4]Q. Wang, Y. Hasegawa, and T. A. Zaki, "Spatial reconstruction of steady scalar sources from remote measurements in turbulent flow," J. Fluid Mech. **870**, 316–352 (2019).
[5]P. Matharu and B. Protas, "Optimal closures in a simple model for turbulent flows," SIAM J. Sci. Comput. **42**, B250–B272 (2019).
[6]Y. Li, J. Zhang, G. Dong, and N. S. Abdullah, "Small-scale reconstruction in three-dimensional Kolmogorov flows using four-dimensional variational data assimilation," J. Fluid Mech. **885**, A9 (2020).
[7]P. Bauweraerts and J. Meyers, "Reconstruction of turbulent flow fields from lidar measurements using large-eddy simulation," J. Fluid Mech. **906**, A17 (2020).
[8]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).
[9]C. He, Y. Liu, and L. Gan, "Instantaneous pressure determination from unsteady velocity fields using adjoint-based sequential data assimilation," Phys. Fluids **32**, 035101 (2020).
[10]K. Duraisamy, G. Iaccarino, and H. Xiao, "Turbulence modeling in the age of data," Annu. Rev. Fluid Mech. **51**, 357–323 (2019).
[11]K. Yoshida, J. Yamaguchi, and Y. Kaneda, "Regeneration of small eddies by data assimilation in turbulence," Phys. Rev. Lett. **94**, 014501 (2005).
[12]C. C. Lalescu, C. Meneveau, and G. L. Eyink, "Synchronization of chaos in fully developed turbulence," Phys. Rev. Lett. **110**, 084102 (2013).
[13]P. C. D. Leoni, A. Mazzino, and L. Biferale, "Inferring flow parameters and turbulent configuration with physics-informed data assimilation and spectral nudging," Phys. Rev. Fluids **3**, 104604 (2018).
[14]P. C. D. Leoni, A. Mazzino, and L. Biferale, "Synchronization to big data: Nudging the Navier-Stokes equations for data assimilation of turbulent flows," Phys. Rev. X **10**, 011023 (2020).
[15]A. Vela-Martin, "The synchronisation of intense vorticity in isotropic turbulence," J. Fluid Mech. **913**, R8 (2021).
[16]M.-A. Nikolaidis and P. J. Ioannou, "Synchronization of low Reynolds number plane Couette turbulence," J. Fluid Mech. **933**, A5 (2022).
[17]M. Wang and T. A. Zaki, "Synchronization of turbulence in channel flow," arXiv:2111.12004 [physics.flu-dyn] (2021).
[18]D. P. G. Foures, N. Dovetta, D. Sipp, and P. J. Schmid, "A data-assimilation method for Reynolds-averaged Navier-Stokes-driven mean flow reconstruction," J. Fluid Mech. **759**, 404–431 (2014).
[19]L. M. Pecora and T. L. Carroll, "Synchronization in chaotic systems," Phys. Rev. Lett. **64**, 821–824 (1990).
[20]S. Boccaletti, J. Kurths, G. Osipov, D. L. Valladares, and C. S. Zhou, "The synchronization of chaotic systems," Phys. Rep. **366**, 1–101 (2002).
[21]M. Buzzicotti and P. C. D. Leoni, "Synchronizing subgrid scale models of turbulence to data," Phys. Fluids **32**, 125116 (2020).
[22]M. Zauner, V. Mons, O. Marquet, and B. Leclaire, "Nudging-based data assimilation of the turbulent flow around a square cylinder," J. Fluid Mech. **937**, A38 (2022).
[23]C. Meneveau and J. Katz, "Scale-invariance and turbulence models for large-eddy simulation," Annu. Rev. Fluid Mech. **32**, 1–32 (2000).
[24]S. B. Pope, *Turbulent Flows* (Cambridge University Press, Cambridge, 2000).
[25]C. Rosales and C. Meneveau, "Linear forcing in numerical simulations of isotropic turbulence: Physical space implementations and convergence properties," Phys. Fluids **17**, 095106 (2005).