From limited observations to the state of turbulence: Fundamental difficulties of flow reconstruction

Tamer A. Zaki and Mengze Wang

From limited observations to the state of turbulence:
Fundamental difficulties of flow reconstruction

Tamer A. Zaki[*] and Mengze Wang

Department of Mechanical Engineering, Johns Hopkins University, Baltimore, MD 21218, USA

Numerical simulations of turbulence provide non-intrusive access to all the resolved scales and any quantity of interest, although they often invoke idealizations and assumptions that can compromise realism. In contrast, experimental measurements probe the true flow with lesser idealizations, but they continually contend with spatio-temporal sensor resolution. Assimilating observations directly in simulations can combine the benefits of both approaches and mitigate their respective deficiencies. The problem is expressed in variational form, where we seek the flow field that satisfies the Navier-Stokes equations and minimizes a cost function defined in terms of the deviation of the computational predictions and available observations. In this framework, measurements are no longer a mere record of the instantaneous, local quantity, but rather an encoding of the antecedent flow events that we aim to decode using the governing equations. Chaos plays a central role in obfuscating the interpretation of the data: observations that are infinitesimally close may be due to entirely different earlier conditions. We examine a number of state estimation problems: In circular Couette flow, starting from observations of the wall stress, we accurately reconstruct the wavy Taylor vortices that interact nonlinear to maintain a saturated state. Through a discussion of transition to chaos in a Lorenz system, we highlight the challenge of navigating the landscape of the cost function, and how the landscape can favorably be modified by sensor weighting and placement. In turbulent channel flow, the Taylor microscale and Lyapunov timescale place restrictions on the resolution of observations for which we can accurately reconstruct all the missing scales. The notion of domain of dependence of an observation is introduced and related to the Hessian of the cost function. For measurements of wall shear stress, the eigenspectrum of the Hessian demonstrates the sensitivity of short-time observations to the fine near-wall turbulent scales and to the large scales only in the outer flow. At long times, backward chaos obfuscates the interpretation of the data: observations that are infinitesimally close may be traced back to entirely different earlier flow states—a dual to the more common butterfly effect for forward trajectories.

I. INTRODUCTION

A. Why data assimilation?

Flow simulations are continually pushing computational boundaries to establish new frontiers, for example turbulence at progressively higher Reynolds numbers [1, 2] and phenomena that take place at hypersonic speeds [3]. Simulations also attempt to probe ever more complex configurations that involve multiphase, rheological, physiological effects [4–7]. Despite impressive advancements, even at the highest levels of simulation fidelity some modeling assumptions invariably remain, including treatment of boundary conditions such as periodic and truncated simulation domains [8, 9] and initial conditions when multiple flow states exist for the same set of parameters [10]. At extreme conditions, models are adopted to reduce the computational cost or to account for unknown physics. Choices of model parameters, for example in sub-grid stress and wall models in large-eddy simulations of high Reynolds number or in non-equilibrium chemical models at high speeds, have direct impact on the accuracy and robustness of simulations.

Experiments on the other hand provide a level of realism that is difficult to reproduce computationally, and can access extreme flow regimes, including very high Reynolds numbers that remain, to date, intractable computationally [11]. Experiments must, however, contend with the difficulty of performing non-intrusive measurements and the limitations of sensor resolution. In some instances, it is not possible to directly measure a quantity of interest and it can only be inferred or computed from other measurements. For example, while we may be interested in the pressure field in a turbulent flow, the experimental measurements may only provide successive images of particle tracers that must be interpreted and processed to compute the velocity field and subsequently the associated pressure [12].

---

[*] corresponding author, email: t.zaki@jhu.edu

Data assimilation combines the respective strengths of both experiments and simulations, and mitigates their deficiencies [13, 14]: By infusing experimental measurements in simulations, we enhance the fidelity of the computations and mitigate epistemic uncertainties. At the same time, the numerical prediction are not only consistent with the observations, they also provide non-intrusive access to the full flow state at much higher resolution than the original data [15]. The benefit of solving these inverse problems, from observations to flow state, extends beyond simply enhancing the resolution of scarce data. Inverse problems provide a unique perspective for the interpretation of observations. A sensor signal is no longer a mere record of a measured quantity at a point in space and an instance in time, but rather an encoding of the antecedent flow events that led to the measurement [16]. By decoding the observations we reveal the hidden flow physics that is unavailable from forward analysis, including the domain of dependence of an observation site, the critical data resolution for an accurate state estimation, and the robustness of reconstructed flow events to observation noise. This perspective also motivates a new notion of optimal layout and weighting of measurements: Instead of concentrating measurements near a region of interest based on knowledge of the forward problem, sensor placements can be optimized in order to provide the most accurate interpretation, or reconstruction, of the entire flow field using data assimilation [17, 18].

B. Choice of approach

In choosing a data-assimilation approach for the interpretation, or augmentation, of observations, it is important to differentiate some of the existing classes of techniques. One distinction is among methods that actually satisfy the governing flow equations versus those that approximate the velocity field using a statistical or reduced-order, perhaps machine-learned, model [19–22]. Here our focus will be on predictions that satisfy the full Navier-Stokes equations and that best reproduce observations. This choice is desirable if one is interested in the fundamental flow dynamics, as we are here, or in forecasting beyond the observation time horizon.

Another classification is based on how observations are assimilated into numerical simulations. Filters use the latest observations to update the state [23, 24], while smoothers attempt to identify the state whose evolution reproduces all the observations within the entire assimilation horizon [14, 25]. Again, since our interest is in the evolution of the flow state, or the dynamics, we will focus on smoothers. In this context, we define a cost function which is the difference between model predictions and observations, and we proceed to identify the flow state that (i) satisfies the governing Navier-Stokes equations and (ii) minimizes the deviation from observations. The minimization procedure requires the gradient of the cost function with respect to the initial flow state; this gradient also quantifies the sensitivity of observations to initial flow field.

Two smoothing techniques have been adopted successfully in turbulent flows: adjoint- and ensemble-variational methods. Adjoint methods are the most efficient in terms of evaluating the gradient [26]. The associated computational cost is always on the order of twice that of calculating the cost function, regardless of the size of the control vector. However, for chaotic systems the accuracy of the gradient deteriorates for long integration time horizons [27, 28]. Therefore, the implementation of adjoint methods is restricted to short assimilation windows that can be concatenated to span long periods of observations [20, 29]. Ensemble methods do not require an adjoint model [17]. They use an ensemble of estimates of the state and advance them in time, collect model predictions and compare them to observations. The local gradient of the cost function with respect to the state can be approximated at the location of the mean of the ensemble using a quadratic approximation, and used to update the estimated state. These methods are suitable for long horizons, for example when using statistical observations. The number of ensemble members, however, increases with the size of the control vector and they are not necessarily as robust as adjoint techniques. Some methods combine the idea of ensemble and adjoint to evaluate the sensitivity of statistical observations to model parameters [28]. Here we adopt adjoint-variational method, because we are interested in fundamental questions regarding the domain of dependence of observations, the minimum observation resolution that enable reconstruction of the entire turbulent field, and the robustness of reconstruction to observation noise.

In the next section, we introduce the adjoint-variational approach briefly. We proceed to apply the method to the interpretation of measurements in circular Couette flow. The configuration is designed such that many saturated states are possible, and we attempt to identify the particular one that generated the wall stresses. In this case, the dynamics are not chaotic. We proceed to discuss what to expect in case of transition to turbulence, and briefly remark on some recent results in that domain. We then turn to canonical turbulent channel flow, explore the critical data resolutions that enable an accurate turbulence reconstruction, and discuss the notion of domain of dependence of measurements. Concluding remarks are provided in the final section.

II.   ADJOINT VARIATIONAL DATA ASSIMILATION

Since our interest is in a high-fidelity interpretation of flow observations, throughout we will adopt the Navier-Stokes equations to predict the evolution of the flow and will resolve all the flow scale using direct numerical simulations. This choice eliminates any turbulence modeling assumptions. Our flow thus satisfies the incompressibility constraint and momentum equations,

$$\nabla \cdot \boldsymbol{u} = 0 \tag{1}$$

$$\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u} \cdot \nabla \boldsymbol{u} = -\nabla p + \frac{1}{Re} \nabla^2 \boldsymbol{u}. \tag{2}$$

These equations are also referred to as the forward model in adjoint-variational method, and the velocity $\boldsymbol{u}$ and pressure p as the forward state variables.

In lieu of experimental measurements, we perform independent simulations from a reference initial condition $\boldsymbol{u}_0^R$ and extract surrogate observations,

$$\boldsymbol{m}_n = \mathcal{M}(\boldsymbol{u}_n^R). \tag{3}$$

The reference initial state $\boldsymbol{u}_0^R$ then becomes the hidden truth that we aim to discover from its observations $\boldsymbol{m}_n$ alone. The objective of adjoint-variational data assimilation is to identify an estimated initial condition $\boldsymbol{u}_0$ that reproduces the observations, by minimizing a cost function that is proportional to the difference between the observations and their estimation,

$$J(\boldsymbol{u}_0) = \sum_{n=0}^{N} \frac{1}{2} \|\boldsymbol{m}_n - \mathcal{M}(\boldsymbol{u}_n)\|_O^2. \tag{4}$$

The norm $\|\cdot\|_O$ represents evaluation over the observation space, and $\boldsymbol{u}_n$ is the Navier-Stokes evolution of the estimated initial condition. The estimated initial condition is sought iteratively using a gradient-based minimization of the cost function.

In order to evaluate the gradient of J, we introduce the Lagrangian

$$L = J - \left[\boldsymbol{u}^\dagger, \frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u} \cdot \nabla \boldsymbol{u} + \nabla p - \frac{1}{Re} \nabla^2 \boldsymbol{u} \right] - \left[p^\dagger, \nabla \cdot \boldsymbol{u} \right], \tag{5}$$

where the multipliers $\boldsymbol{u}^\dagger$ and $p^\dagger$ are the adjoint velocity and adjoint pressure, respectively. The inner product is an integration over the spatial domain and the time horizon,

$$[\boldsymbol{f}, \boldsymbol{g}] = \int_0^T \int_V \boldsymbol{f}^\top \boldsymbol{g} \ d^3\boldsymbol{x} dt, \tag{6}$$

where $t = 0$ and $t = T$ represent the start and end of the assimilation window. First-order optimality requires that the derivatives of L with respect to both the adjoint and forward variables be zero. The former condition yields the Navier-Stokes equations (1-2) which are enforced by simulating the forward evolution of the estimated field. The latter condition, namely that the derivatives of L with respect to $\boldsymbol{u}_n$ vanishes, yields the adjoint equations,

$$\nabla \cdot \boldsymbol{u}^\dagger = -\frac{\partial J}{\partial p} \tag{7}$$

$$\frac{\partial \boldsymbol{u}^\dagger}{\partial \tau} - \boldsymbol{u} \cdot \nabla \boldsymbol{u}^\dagger + (\nabla \boldsymbol{u}) \cdot \boldsymbol{u}^\dagger = \nabla p^\dagger + \frac{1}{Re} \nabla^2 \boldsymbol{u}^\dagger + \frac{\partial J}{\partial \boldsymbol{u}}, \tag{8}$$

where $\tau \equiv T - t$ is the reverse time. The adjoint system is driven by the source term, which is derived analytically from the cost function (4) and determined by the mismatch between the observations and their estimation, $\mathcal{E}_n = \mathcal{M}(\boldsymbol{u}_n) - \boldsymbol{m}_n$.

The gradient of the cost function is determined by the adjoint field at the initial time $t = 0$ ($\tau = T$),

$$\nabla_{\boldsymbol{u}_0} J = \frac{\partial L}{\partial \boldsymbol{u}_0} = \boldsymbol{u}_0^\dagger. \tag{9}$$

Therefore, the adjoint field $\boldsymbol{u}_0^\dagger$ is the sensitivity of the cost function to initial conditions. This sensitivity can then be used in gradient-based optimization algorithms to minimize J. In the examples presented in this paper, we adopt the limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) method [30]. The above presentation was in terms of
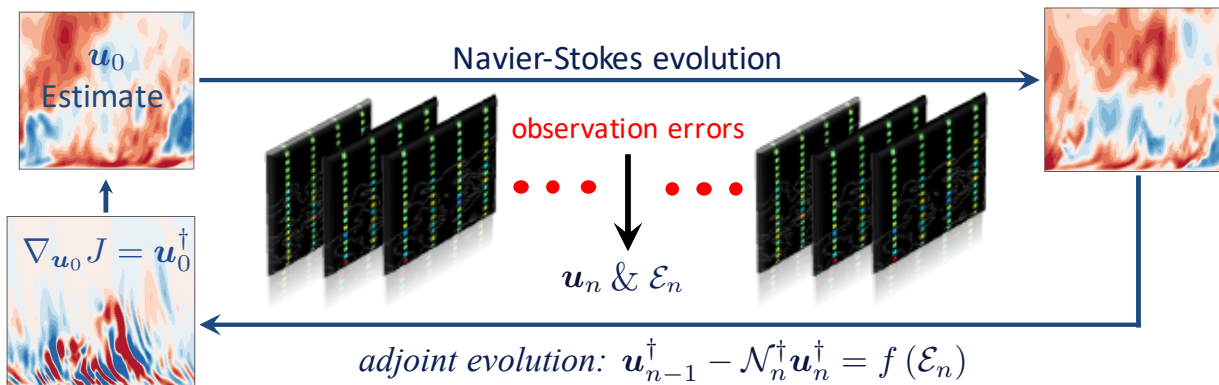
FIG. 1. Schematic of adjoint-variational data assimilation. Starting from an estimate of the initial condition $\boldsymbol{u}_0$, we compute the Navier-Stokes evolution. The forward fields $\boldsymbol{u}_n$ and the deviation from observations $\mathcal{E}_n$ are stored. The adjoint equations are evolved from the final to the initial time, driven by observations errors. The initial adjoint field $\boldsymbol{u}_0^\dagger$ is the gradient of the cost function, and is used to update $\boldsymbol{u}_0$. The forward-adjoint loop is repeated until convergence.

the continuous form of the equations. Discretization of the continuous forward and adjoint systems may affect the accuracy of evaluating the gradient of the cost function. To obtain the exact gradient, we adopted the discrete adjoint approach where we first discretize the forward Navier-Stokes equations and then derive the discrete adjoint system.

The adjoint-variational data assimilation procedure is presented schematically in figure 1. Starting from an estimate of the initial flow state $\boldsymbol{u}_0$, the Navier-Stokes equations are evolved over the assimilation time horizon. During the forward evolution, both the velocity fields and the discrepancies between the observations and their estimates are stored. At the end of the assimilation horizon, the adjoint equations are advanced backward in time, forced by the errors in the observations. The resulting gradient of the cost function at the end of adjoint marching is used in the L-BFGS algorithm to update the estimated initial condition. This procedure is then repeated for a prescribed number of iterations or until a convergence criterion is reached.

The utility of data assimilation is far reaching, beyond the important application of augmenting experimental measurements. These techniques enable us to pose new questions. In turbulence, for instance, we can progressively increase the sparsity of observations and ask what is the largest spatio-temporal spacing that still enables full reconstruction of the missing scales, and why? The adjoint-variational approach in particular provides a unique perspective for interpretation of the flow physics. Adjoint fields quantify the sensitivity of observations to the flow state. When observations are spatio-temporally distributed, comparing their adjoint fields provides guidance for sensor placement and relative weighting, and motivates new strategies to optimize these sensing parameters. Furthermore, adjoint equations can be derived to evaluate the domain of dependence of an isolated measurement and the Hessian of the cost function at optimality, which yield a unique perspective on the difficulty of estimating the state of turbulence from limited observations. These ideas will be explored in the remaining sections with the aid of specifically designed examples.

III. FLOW RECONSTRUCTION FROM LIMITED MEASUREMENTS

A. Traveling saturated waves in circular Couette flow

The first configuration that we will examine is the circular Couette flow between two concentric cylinders, which was explored in detail by Wang *et al.* [26]. The gap between the inner and outer cylinders is $d = r_o - r_i = 1$, and the ratio of radii is $r_i/r_o = 0.714$. The outer cylinder is stationary, and the inner-cylinder tangential speed v_i is adopted as the reference velocity scale in the definition of the Reynolds number, $Re_i \equiv v_i d/\nu = 400$. The above parameters were selected such that we can demonstrate the method in a flow that involves nonlinearly interacting, unsteady, traveling waves, yet the total energy is constant and the flow is not chaotic. Another important feature is that for the same Reynolds number, there are multiple admissible flow states and which one is observed depends on the precise experimental geometry and protocol, for example the aspect ratio of the cell, the initial condition and the transients. Simulations may not be able to reproduce all the experimental conditions, which may also include uncertainties. However, using observations from the experiments within the saturated flow regime only, i.e. beyond any experimental transients, we can adopt data assimilation strategies to reproduce the exact experimental state and the precise evolution of the constituent waves as a function of time.
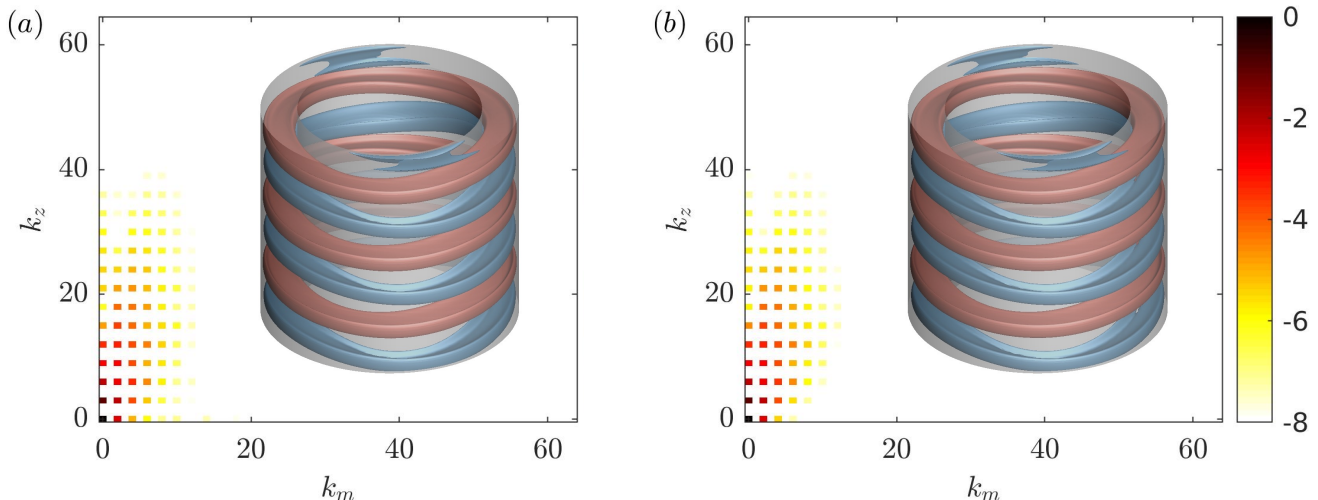
FIG. 2. Energy spectra averaged across the gap, $\log_{10}\langle\|\hat{\boldsymbol{u}}\|^2\rangle_d$, where $\hat{\boldsymbol{u}}$ is the Fourier transform of the velocity perturbation field $\boldsymbol{u} - \langle\boldsymbol{u}^R\rangle$. Insets: isosurfaces of (red) positive and (blue) negative azimuthal velocity fluctuations, $v_\theta - \langle v_\theta^R\rangle = \{0.14, -0.14\}$. (a) Adjoint-variational estimation at $t = T$; (b) the true state. Both spectra are normalized by the maximum value of the true spectrum.

The observations that we attempt to interpret are the shear stresses on the cylinder walls as a function of time, which are imprinted by the signature of the nonlinearly interacting, traveling, wavy Taylor rolls. These observations spanned a time horizon $T = 20$ time units, which is more than one rotation of the inner cylinder, and were collected from an independent computation. The associated flow is the hidden truth that we aim to discover.

Our first estimate of the flow state can simply be the laminar profile between two concentric cylinders. This state naturally fails to reproduce the observations since it is void of any Taylor vortices. At the end of the first forward simulation, the adjoint model is then forced by the discrepancy in the wall shear stresses from the first guess and available observations. Upon reaching the initial time, the adjoint field provides the gradient of the cost function with respect to the initial condition, which is used in the L-BFGS algorithm to obtain the update direction of the initial flow state. This procedure is repeated iteratively until our computational predictions reproduce the available data of the wall stresses to within a prescribed level of accuracy.

Results reported here were obtained after one hundred forward-adjoint loops in order to ensure accuracy. The resulting decrease in the cost function was three orders of magnitude from its initial value. The predicted state $\boldsymbol{u}$ at the end of the assimilation time horizon is visualized in figure 2 where it is compared to the hidden truth $\boldsymbol{u}^R$. The iso-surfaces show positive and negative tangential velocity perturbations. Qualitatively, the estimated flow has the correct number of wavy Taylor rolls and azimuthal phase of the structures. Quantitatively, the root-mean-squared error $\mathcal{E}_V = \langle\|\boldsymbol{u} - \boldsymbol{u}^R\|^2\rangle_V^{1/2}$, averaged over the volume V, is less than 5%. The spectra of the assimilated and true states are also compared at the final time. This plot highlights the presence of a large number of axi-symmetric and wavy vortices, traveling around the azimuth and interacting nonlinearly. The agreement in the spectra between the assimilated and true states is evident, in particular for energetic modes that represent dominant flow structures. The scale in the figure is selected to highlight the accuracy of the prediction across eight orders of magnitude.

This first example is by no means trivial as can be inferred form the energy spectra. The results demonstrate that data assimilation can be adopted to augment limited experimental observations, and mitigate uncertainty in computational parameters. Specifically, the adjoint approach did not require any knowledge of the startup procedure of the true flow that led to the saturated state when the measurements were collected. Instead, the forward-adjoint loop relied on the measurements to directly discover the initial condition within the energy saturated state and its time evolution. It should be noted, however, that this first example was designed to eliminate a challenge, namely estimation of chaotic dynamics. Specifically, in the present case, once the energy of the system reaches the saturated state, the flow is stable. What happens when the flow is chaotic? Before tackling the problem of estimating a fully turbulent state, it is instructive to examine a transitional system where organized and chaotic dynamics are juxtaposed.
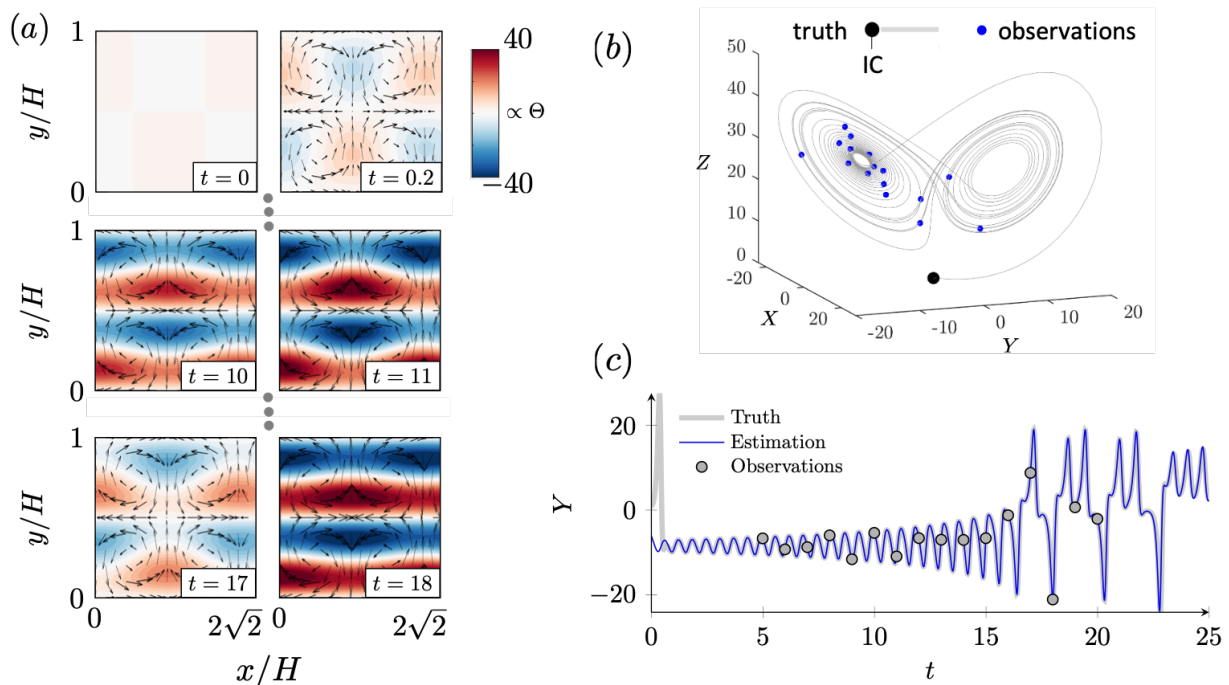
FIG. 3. (a) Evolution of the Lorenz system in physical space. Contours are the temperature perturbations and arrows show the velocity field. (b) State-space evolution of the Fourier coefficients, starting from the true initial condition with marked observations. (c) Time evolution of Y from the (gray) true and (blue) estimated initial conditions, with marked observations. The assimilated trajectory latches onto the truth, reproduces observations, and accurately forecasts the system for some time beyond the observation horizon [18].

B. Transition to chaos

Laminar-to-turbulence transition is difficult to predict due to its chaotic nature. Infinitesimal changes in the system state can lead to significant modification of transition characteristics. For example, spontaneously formed turbulent spots in bypass transition of boundary layers [31–33] and puffs in pipe flows [34, 35] appear sporadically and are difficult to anticipate. As the system undergoes transition from an organized to a chaotic state, it is intuitive to expect that early observations are more valuable, or beneficial, when attempting to estimate the initial condition. A cost function J defined using sensors in the laminar region will vary smoothly when estimates of the initial state approach the truth, which can promote convergence of the optimization procedure. In contrast, when sensors are in the turbulent region, the cost function oscillates significantly in the vicinity of the measurements since small changes in the initial condition can lead to significantly different observations and, as a result, convergence may be more difficult.

A helpful prototype system to explore the influence of transition and chaos on data assimilation is the Lorenz model of unstable convection in two dimensions [18]. The setup is periodic in the horizontal direction, and both gravity and a temperature difference between the lower and upper boundaries, $\Delta\Theta = \Theta(0) - \Theta(H)$, act in the vertical direction. Small initial perturbations amplify in time, and lead to convection cells that oscillate about an equilibrium and suddenly overturn (figure 3). The process of oscillation and overturning repeats in a chaotic fashion [36]. The early amplification of perturbations is akin to growth of instability waves in wall-bounded flows, and the resulting sporadic overturning of the cells is a surrogate for the chaotic inception of turbulence.

For the Lorenz system, the temperature fluctuations Θ and velocity potential ψ can be expressed in Fourier space in terms of the horizontal and vertical wavenumbers $k_a = a2\pi/H$ and $k_y = 2\pi/H$,

$$\Theta(x,y,t) \propto Y(t)\sqrt{2}\cos\left(\tfrac{1}{2}k_a x\right)\sin\left(\tfrac{1}{2}k_y y\right) - Z(t)\sin\left(k_y y\right) \tag{10}$$

$$\psi(x,y,t) \propto X(t)\sqrt{2}\sin\left(\tfrac{1}{2}k_a x\right)\sin\left(\tfrac{1}{2}k_y y\right). \tag{11}$$

The time-evolution of the Fourier coefficients $\boldsymbol{Q} = [X,Y,Z]^{\top}$ is given by,

$$\frac{dX}{dt} = \sigma(Y - X), \quad \frac{dY}{dt} = -XZ + \rho X - Y, \quad \text{and} \quad \frac{dZ}{dt} = XY - \beta Z. \tag{12}$$
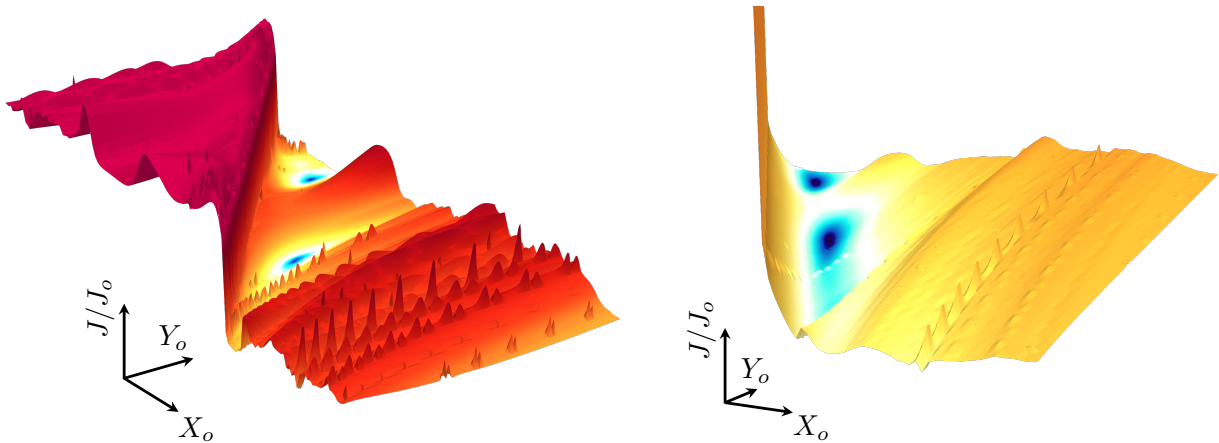
FIG. 4. The dependence of the cost function of the Lorenz system on (X_o, Y_o) near the global optimal. Left: The landscape of the cost function when all observations are uniformly weighted. Right: The landscape of J when the observations are weighted by a Gaussian that preferentially weights early measurements.

Similar to Saltzman [37], we adopt the initial condition $[X_o, Y_o, Z_o]^\top = [0, 1, 0]^\top$ which becomes our hidden truth, and we consider $a = \sqrt{1/2}$, $\sigma = 10$, $\rho = 28$. The small size of the system enables us to efficiently evaluate and visualize the cost function, and to explore ideas related to sensor placement and weighting [see also appendix by 18].

Consider when observations are available for Y only, at time instances $t_m = \{5, 6, \ldots, 20\}$ which span the early periodic solution and later chaotic behavior (see figure 3). We attempt to discover the initial condition $[X_o, Y_o, Z_o]^\top$ that reproduces the measurement vector $\boldsymbol{m} = [Y(t_m)]^\top$ by minimizing the cost function $J = \frac{1}{2}\|\boldsymbol{m} - \mathcal{M}(Y)\|^2$. For this small system, we can directly evaluate J and plot its dependence on the control vector near our estimate. Figure 4 shows that the cost-function landscape is highly oscillatory, laden with local minima and, as a result, difficult to navigate to identify the global minimum. The idea of preferentially weighting early observation was evaluated by Buchta and Zaki [18]. The weighted cost function is $J = \frac{1}{2}\|\boldsymbol{W}(\boldsymbol{m} - \mathcal{M}(Y))\|^2$, where $\boldsymbol{W}$ is a diagonal matrix with larger elements corresponding to earlier observations. The resulting cost-function landscape (figure 4) is much smoother and amenable to efficient gradient-based optimization.

Using the preferential weighting of upstream observations, [18] performed an ensemble, rather than adjoint, variational estimation of the initial condition and the results are shown in figure 3c. While the predicted initial state is visibly displaced relative to the truth, the predicted trajectory quickly latches onto the true one, reproduces all the observations and provides an accurate forecast for the shown duration past the observation horizon.

While the idea of sensor weighting was motivated here by physical arguments, it is based on rigorous mathematical foundation. The shape of the cost function depends on the observation operator $\mathcal{M}$, which determines the measured quantities and their placements, and the measurement weighting $\boldsymbol{W}$. Since at optimality the gradient of J vanishes, we are particularly interested in how the design of observations affects the Hessian of the cost function. Preferentially weighting early measurements reduces the maximum eigenvalue of the Hessian, thus reducing extreme curvature. Buchta and Zaki [18] adopted this strategy not only for the Lorenz system above, but also for the much more complex and higher dimensional configuration a transitional, high-speed boundary layer. They were able to achieve significant improvement in the accuracy of inflow estimation from discrete wall-pressure probes by preferentially weighting sensors upstream of transition. It should be cautioned, however, that weighting early observations should not be naively adopted. In a transitional boundary layer, for instance, a wall-pressure probe should not be placed arbitrarily upstream. An obvious consideration is the potentially low signal-to-noise ratio. Additionally, if placed ahead of the receptivity site or, more subtly, in a region where the sensor does not have sensitivity to the incoming instability wave, the associated observations can not assist in minimizing the cost function. Also note that, in a fully turbulent flow, the cost landscape is mountainous independent of the choice of observed quantities, their spatio-temporal placement and their relative weighting. In that case, optimal sensor placement and weighting can target improving the condition number of the Hessian, which has been adopted with success in problems involving prediction of scalar sources in turbulent environments [16, 17]. We will return to a discussion of the Hessian after we examine the accuracy of adjoint-variational estimation of channel-flow turbulence from limited observations.

As we turn to the fully turbulent case, we should keep in mind important points that are well captured by the results of the Lorenz system. Firstly, due to the intricate landscape and numerous local minima of the cost function (figure 4), the first guess, or estimate, of the state in the variational algorithm has an impact on convergence and, hence, the final estimated state. Therefore, the construction of the first guess should exploit any known characteristics
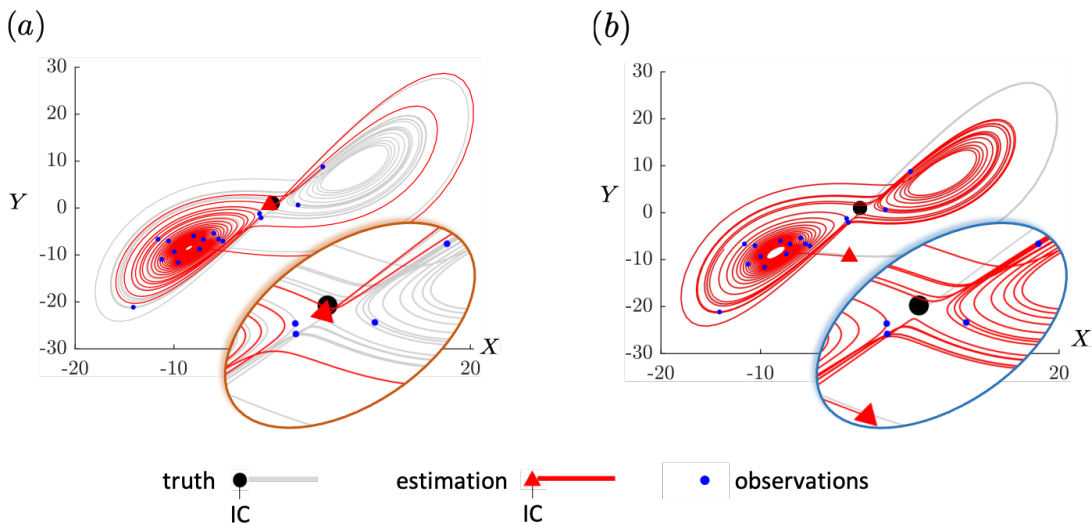
FIG. 5. State-space evolution from estimates of the initial Fourier coefficients of the Lorenz system. (a) Divergence of trajectories starting from the true and a slightly perturbed initial condition. (b) Trajectory from estimated initial condition. Despite a large initial separation from the truth, the assimilated trajectory latches onto the truth, reproduces observations, and accurately forecasts the system for some time beyond the observations horizon.

of the initial state. Absent such knowledge, multiple starting points for the assimilation procedure can be adopted. Secondly, for chaotic nonlinear systems, there is no guarantee of a unique solution to the inverse problem. The results in figure 3c show that measurements, and extended portions of the trajectory, can be accurately reproduced from more than one initial condition. This point can be further clarified by contrasting the "butterfly effect" and its dual, or "adjoint butterfly effect", which are illustrated in figure 5. In the former case, a small change in the initial condition leads to significant deviation in the forward evolution due to chaos, and the same is of course true of turbulence. In practical terms, a simulation that starts from a perfectly measured state will diverge from the experiment and from other simulations that have as little as different round-off errors. This behavior can impede our attempt to predict the initial condition that reproduces the measurements, limit the assimilation time horizon and compromise the accuracy of forecasts. As for the dual butterfly effect, it may be less familiar but is crucially important, especially for the adjoint approach. Infinitesimally close observations can be reproduced by initial conditions that have large separation as we remarked earlier. In terms of the adjoint, the small deviations in the observations lead to entirely different adjoint evolution, and are traced back to different initial conditions. From a practical standpoint, errors in the observations can thwart accurate estimation of the state.

C. Turbulent channel flow

In this section, we will attempt to reconstruct fully turbulent channel flow from sparse velocity data, and explore the critical data resolution that guarantees an accurate estimation. In order to minimize uncertainties, have full control of the resolution of observations and be able to assess the accuracy of our reconstruction, we extracted our observations from an independent direct numerical simulation (DNS). The flow in the channel was driven by a fixed pressure gradient at $Re_\tau = u_\tau h/\nu = 180$, where u_τ is the friction velocity and h is half-channel height. The observations are collected by sub-sampling the velocity field as shown in figure 6a. The observations were distributed in the streamwise, wall-normal and spanwise directions separated by $\Delta x_m^+ = 47$, $\Delta y_m^+ \in [1.8, 24]$ and $\Delta z_m^+ = 28$, respectively, and their temporal resolution was $\delta t_m^+ = 0.46$. Taken all together, the observations are $1/4096$ of the data resolution required for direct numerical simulations. The estimation window was $T = 4.2$ ($T^+ = 50$), which was approximately one Lyapunov timescale of the flow [38]. This setup for observations is motivated by the trade-off between field of view and resolution in volume measurements in experiments, for example in particle image velocimetry (PIV).

Estimating the turbulent state by simply interpolating the available sparse observations is fruitless, especially if we are interested in the velocity gradients. Figure 6b shows the outcome of such approach, where only a small portion of the vortical structures are captured in the interpolated flow field. Naturally, the accuracy of the computed structures further deteriorates if the observations are contaminated by noise [39].

Adjoint-variational data assimilation was performed in order to accurately estimate the flow state. The first guess of
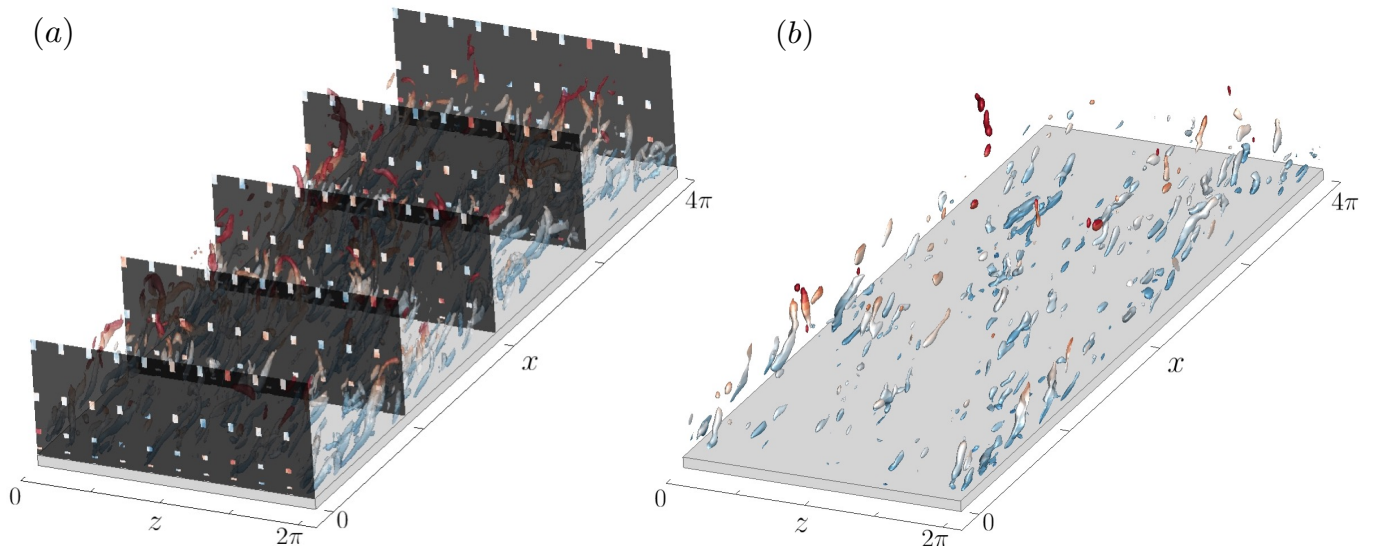
FIG. 6. (a) True vortical structures visualized using the λ_2 vortex identification criterion, with threshold $\lambda_2 = -2$. The isosurfaces are colored by wall-normal height, y. Pixels on the black cross-flow planes represent scarce observations, or sub-sampled velocity data. (b) Vortical structures computed from interpolation of the observations.

the estimated initial condition was a simple interpolation of the observations, and we performed one hundred forward-adjoint loops. The procedure reduced the cost function to 2.8% of its initial value. The accuracy of the predicted flow evolution can be evaluated by comparison to the hidden truth. The volume-averaged estimation error is plotted in figure 7a (black line) as a function of time. Within $t \in [0, T]$ (gray region), the error decays monotonically and the flow more closely shadows the true trajectory in state space. At the end of the assimilation window, $t = T$, the error becomes an order of magnitude smaller than at the initial time. This trend is interesting and can be explained with reference to the notion of adjoint chaos introduced above. Perturbations to the adjoint system amplify in backward time. As such, when one considers the adjoint equations (7-8), the forcing by the cost function at final time has the maximum potential for amplification during the backward evolution to the initial condition. More precisely, the disparity between model predictions and observations at late times have the most pronounced impact on the gradient direction for updating the initial condition. As a result, while the assimilated initial state is thus optimized to reproduce the time-history of observations, it places progressively higher emphasis on later ones.

As commonly known based on everyday experience with weather prediction, forecasts are progressively less accurate. The same is true here: beyond the observation horizon, the errors in the estimated flow amplify although accuracy within $[T, 2T]$ may be considered acceptable since it is commensurate with the levels during the assimilation horizon. It is important to note that the reported levels of error are on the order of a percent. While the levels of errors are much more restrictive than the public demands for weather forecasts, they are essential for accurate interpretation of flow experiments and probing the flow physics. The errors amplify exponentially due to the chaotic nature of the flow and finally saturate when the turbulence in the channel is entirely decorrelated from the true flow. Should new observations become available, adjoint-variational data assimilation can be performed again starting from the previous estimate of the flow. In figure 7a, new observations were made available within $t \in [2T, 3T]$ (light green region). The data assimilation procedures were repeated, and was able to drive the estimated trajectory towards the true state, again (green line). This figure nicely captures the influence of both forward and adjoint chaos: The very small differences between the green and black curves at $t = T$ lead to significant differences at long times due to forward amplification of perturbations. Meanwhile, the errors in the green curve decay during the assimilation horizon $t \in [2T, 3T]$ due to amplification of perturbations in backward time during the adjoint estimation.

The robustness of the algorithm for predicting vorticity from noisy observations is examined in figure 7b. For references, we are also plotting the errors from simple interpolation of the noisy observations (dashed curve) when the noise level is 10%. In this case, the errors in the near wall region reach approximately 30% of the mean wall vorticity. By contrast, the adjoint-variational estimation (black solid line) reduces the error to 4% of the wall vorticity. The reconstructed vortical structures are visualized in figure 7c and compared to the true flow in figure 7d. Although some of the small-scale structures are not perfectly captured, most of the estimated wall-attached and detached vortical structures are almost identical to the true flow. This compelling estimation accuracy demonstrates the capacity of our algorithm to augment under-resolved turbulence data.
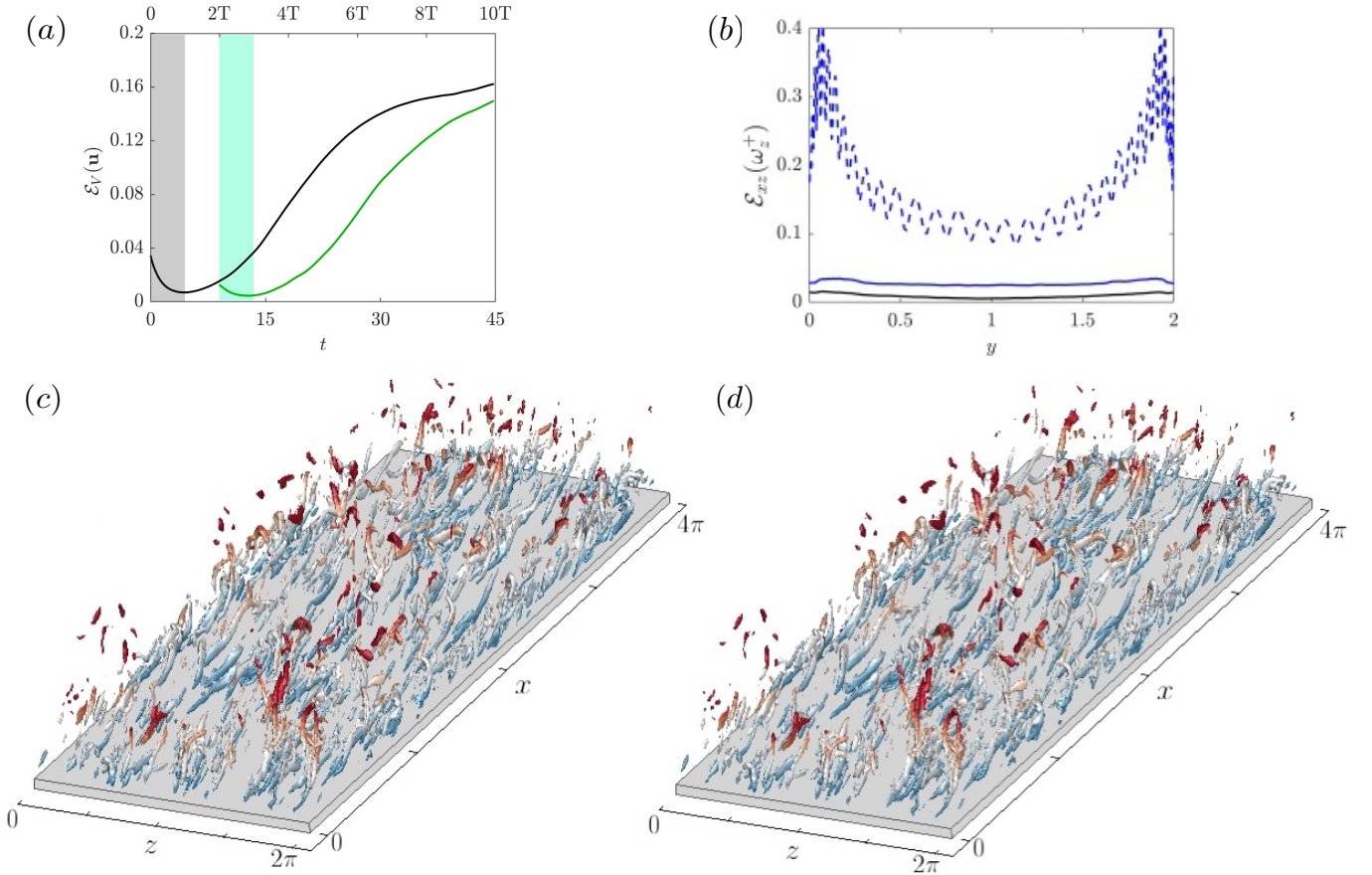
FIG. 7. (a) Temporal evolution of volume-averaged error of the adjoint-variational estimated velocity field. Black line: initial condition at $t = 0$ is estimated using data within $t \in [0, T]$; Green line: flow state at $t = 2T$ is estimated using data within $t \in [2T, 3T]$. (b) Horizontally averaged error $\mathcal{E}_{xz}(\omega_z^+)$ in spanwise vorticity field at $t = T$. (Solid) Adjoint-variational estimation was performed using velocity data that is (black) noise-free and (blue) contaminated by 10% noise relative to the local velocity. (Blue dashed) Vorticity is evaluated directly from interpolation of noisy velocity data. Errors are normalized by mean vorticity $\langle \partial u / \partial y \rangle$ at the wall. (c,d) Estimated and true vortical structures at $t = T$, visualized using the λ_2 vortex identification criterion, with threshold $\lambda_2 = -2$ and colored by wall-normal distance y.

Intuitively, as the spatio-temporal spacing of observations is increased, reconstruction of the entire turbulent state becomes more difficult. Through a systematic effort and numerous data assimilation experiments, we explored the critical resolution of observations that can enable an accurate reconstruction of the turbulence, with 90% correlation coefficient between the assimilated and true states. Two important physical scales emerged as key factors: the Taylor microscale Λ and the Lyapunov timescale τ_σ. The former is evaluated from the two-point velocity correlation, but perhaps more informative is its physical interpretation in the context of isotropic turbulence where it measures the distance traversed by a Kolmogorov eddy during its lifetime as it is swept by the root-mean-squared velocity. This description can be reframed in backward time: An observation of a Kolmogorov eddy will be swept that distance during its adjoint lifetime. As a result, if observations are spaced at twice the Taylor microscale, their signal can be decoded to accurately reconstruct the entire volume. The Lyapunov timescale τ_σ is the e-folding time of an initial disturbance. Even if the observation sites are closer than 2Λ, the temporal sampling interval cannot exceed τ_σ or else Lyapunov amplification of errors will become dominant. In effect, the spatio-temporal data resolution must satisfy

$$\Delta_m \lesssim 2\Lambda, \quad \delta t_m \lesssim \tau_\sigma, \tag{13}$$

to guarantee an accurate turbulence reconstruction. In the direction of mean advection, we can trade spatial and temporal resolutions of observations,

$$\delta t_m \lesssim 2\Lambda/U_a, \quad \Delta x_m \lesssim U_a \tau_\sigma. \tag{14}$$

These criteria were examined in detail and supported by numerical experiments [15] at $Re_\tau = 180$. Future investigations at higher Reynolds numbers, with larger separation of scales, can further verify and refine these conditions.

The criteria discussed above hint to the notion of the domain of dependence of observations. Intuition suggests that this domain depends on the type of observation, its placement in the flow and its separation in time from the initial condition. Absent any advection, we can picture that an isolated observation soon after the initial condition depends on a small spherical region due to diffusion of nearby effects, and that region is larger for later observations. When mean advection is present, we can picture an plume-like region upstream of the observation site, which marks all the potential locations where early events could have reached the sensor. This intuitive picture will be more precise using the adjoint field in the next section.

D. Domain of dependence and Hessian matrix

The domain of dependence of a sensor can be mathematically defined by its sensitivity to the initial flow state. This idea is best explored in terms of variations, specifically whether and to what extent a small change in the initial flow state affects the observation at the sensing location and time. We can therefore consider a small perturbation to the true flow $\boldsymbol{u}' = \boldsymbol{u} - \boldsymbol{u}^R$, which is governed by the linearized Navier-Stokes equations $\boldsymbol{u}' = \mathcal{A}\boldsymbol{u}'_0$, and evaluate its associated influence on the observation, $\mathcal{M}(\boldsymbol{u}) - \mathcal{M}(\boldsymbol{u}^R) = \mathcal{M}(\boldsymbol{u}')$ for a linear observation operator. The same deviation in the observation can also be evaluated using the Lagrangian duality relation,

$$
\begin{aligned}
\mathcal{M}(\boldsymbol{u}') = [\boldsymbol{u}'(t_m), \boldsymbol{\phi}(\boldsymbol{x}_m)] &= [\mathcal{A}\boldsymbol{u}'_0, \boldsymbol{\phi}] \\
&= [\boldsymbol{u}'_0, \mathcal{A}^\dagger\boldsymbol{\phi}] = [\boldsymbol{u}'_0, \boldsymbol{u}^\dagger(\tau = t_m; \boldsymbol{x}_m, t_m)],
\end{aligned}
\tag{15}
$$

where $\boldsymbol{\phi}$ is the observation kernel, and $\mathcal{A}^\dagger$ is the adjoint linearized Navier-Stokes operator. Here we focus on an individual observation from a sensor placed at $\boldsymbol{x}_m$ recorded at the time instant t_m, which is the elementary building block of the cost function. For this instantaneous observation, the inner product in equation 15 only involves spatial integration. As the final expression in equation 15 demonstrates, the sensitivity of observation to the initial state is quantified by the adjoint field $\boldsymbol{u}^\dagger(\tau = t_m; \boldsymbol{x}_m, t_m)$. The support of $\boldsymbol{u}^\dagger$ represents the *domain of dependence* of an observation recorded at $(\boldsymbol{x}_m, t_m)$.

The Lagrangian duality (15) is illustrated in figure 8. In this example, the observation kernel $\boldsymbol{\phi}$ corresponds to evaluating the streamwise wall stress at $(\boldsymbol{x}_m, t_m)$ from the velocity field. The top panel shows the forward approach: we perturb every velocity component at every point in the initial condition separately, and evaluate the corresponding deviation in the wall observation. This approach is therefore computationally expensive because it requires $(nx \times ny \times nz \times 3)$ forward simulations, each evolved until $t = t_m$. The last equality in (15) is significantly more efficient. The adjoint field $\boldsymbol{u}^\dagger(\tau = t_m; \boldsymbol{x}_m, t_m)$ is obtained from a single simulation of the adjoint equations backward in time,

$$
\nabla \cdot \boldsymbol{u}^\dagger = 0
\tag{16}
$$

$$
\frac{\partial \boldsymbol{u}^\dagger}{\partial \tau} - \boldsymbol{u}^R \cdot \nabla \boldsymbol{u}^\dagger + (\nabla \boldsymbol{u}^R) \cdot \boldsymbol{u}^\dagger = \nabla p^\dagger + \frac{1}{Re}\nabla^2 \boldsymbol{u}^\dagger,
\tag{17}
$$

$$
\boldsymbol{u}^\dagger(\tau = 0) = \boldsymbol{\phi}(\boldsymbol{x}_m).
\tag{18}
$$

Note that the adjoint equations (16-18) are different from the earlier adjoint system (7-8) that was driven by a time-dependent forcing from the cost. Here the adjoint field is initiated at $\tau = 0$ from an impulse at the sensor position (Eq. 18), and its evolution uses the true *reference* state $\boldsymbol{u}^R$ (Eq. 17). The resulting adjoint field thus quantifies the sensitivity of wall observations to deviation from the *reference* initial state. Any such deviation $\boldsymbol{u}'_0$ affects the observations at t_m if and only if $\boldsymbol{u}'_0$ is non-zero within the support of $\boldsymbol{u}^\dagger$, as shown in the bottom panel of figure 8.

We will focus the application of the above formulation on wall observations in turbulent channel flow due to their theoretical significance. The generation of all the interior vorticity in the channel can be traced back to the wall [40, 41]. The converse problem, specifically whether the entire initial turbulent state can be decoded from wall signals, has not been addressed comprehensively. Previous efforts using a variety of approaches [15, 19, 42, 43] have demonstrated that the near-wall turbulence and only the outer large-scale structures can be reconstructed from the wall stresses. By quantifying the domain of dependence of wall sensors, we attempt to demystify the difficulty of flow reconstruction from wall observations.

The domain of dependence of a wall stress $\nu\partial u/\partial y$ sensor is in fact the adjoint field in figure 8 (bottom panel). Unlike the forward evolution of an impulse, the adjoint structures are advected upstream of the observation point and expand as a function of reverse time. This behavior was introduced conceptually in the previous section, and is here evaluated using the adjoint equations. The isosurfaces of $u^\dagger$ are symmetric in the spanwise direction because disturbance to the forward u velocity on either side of the measurement cannot be distinguished. At $\tau = t_m$ ($t = 0$), the adjoint field $\boldsymbol{u}^\dagger(\tau = t_m; \boldsymbol{x}_m, t_m)$ measures the sensitivity of wall observation to initial velocity at every point. The
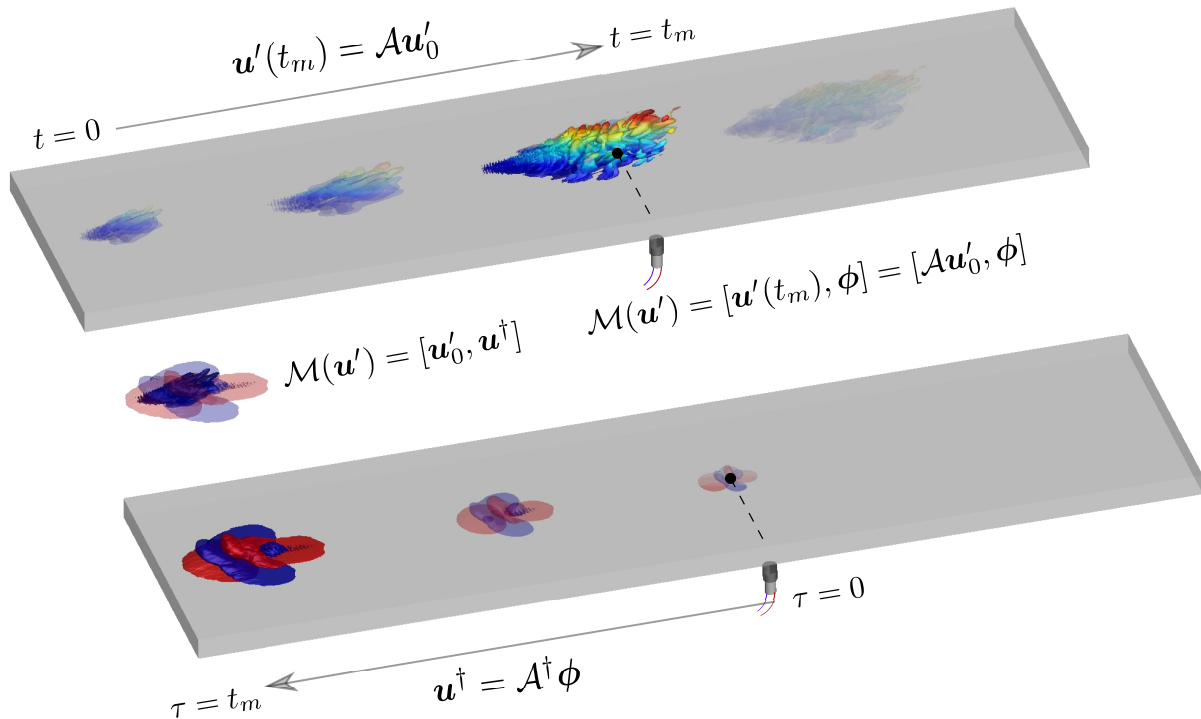
FIG. 8. Schematic of forward-adjoint duality (15) in turbulent channel flow. Top: Forward evolution of an initial perturbation, with isosurfaces showing the streamwise velocity u', colored by wall-normal distance y. The associated impact on the observation is evaluated at the wall at $t^+ = t_m^+ = 20$. Bottom: Backward evolution of the adjoint field starting from an impulse at the wall, up to $\tau = t_m$, where isosurfaces mark (red) positive and (blue) negative adjoint streamwise velocities $u^\dagger$. Modification to the observation due to any initial forward perturbation can be efficiently computed from its inner product with this single adjoint field at $t = 0$ ($\tau = t_m$).

upstream orientation and pancake-like shape of the adjoint isosurfaces in figure 8 implies that the wall observation at $t_m^+ = 20$ is most sensitive to upstream and near-wall perturbations at the initial time.

Since channel-flow turbulence is statistically homogeneous in the horizontal dimensions, observations can be sampled over the entire wall. The cost function associated with small deviation in the observation becomes,

$$\mathcal{J}(\boldsymbol{u}_0'; t_m) = \frac{1}{2S} \int_S \left(\frac{\partial u'}{\partial y}\right)^2 dx_m dz_m = \frac{1}{2S} \int_S [\boldsymbol{u}_0', \boldsymbol{u}^\dagger]^2 dx_m dz_m \tag{19}$$

where the last equality invoked the Lagrangian duality (15) and S is the area of the wall. The cost function (19) is quadratic in $\boldsymbol{u}_0'$, and its gradient vanishes at the true solution ($\boldsymbol{u}_0' = \boldsymbol{0}$). As a result, the difficulty of flow reconstruction near the true solution can be characterized by the Hessian matrix,

$$\mathcal{H}(\boldsymbol{x}_1, \boldsymbol{x}_2; t_m) \equiv \frac{\partial \mathcal{J}}{\partial \boldsymbol{u}_0' \partial \boldsymbol{u}_0'} = \frac{1}{S} \int_S \boldsymbol{u}^\dagger \boldsymbol{u}^\dagger dx_m dz_m. \tag{20}$$

Written explicitly, the Hessian has the form,

$$\mathcal{H}_{ij}(\boldsymbol{x}_1, \boldsymbol{x}_2; t_m) = \frac{1}{S} \int_S u_i^\dagger(\boldsymbol{x}_1, \tau = t_m; \boldsymbol{x}_m, t_m) u_j^\dagger(\boldsymbol{x}_2, \tau = t_m; \boldsymbol{x}_m, t_m) dx_m dz_m, \tag{21}$$

which demonstrates that the Hessian matrix is the auto-correlation of the adjoint field, which corresponds to the observability Gramian in control theory [44]. Here, the Hessian analysis characterizes the difficulty of interpreting observations in order to reconstruct the flow state near optimality, and our particular interest is in non-linear time-dependent turbulent flows.

Eigen-analysis of the Hessian matrix can be viewed from two perspectives: (i) The leading eigenfunctions describe the directions that adjoint-variational method will most effectively target during the data assimilation procedure. In
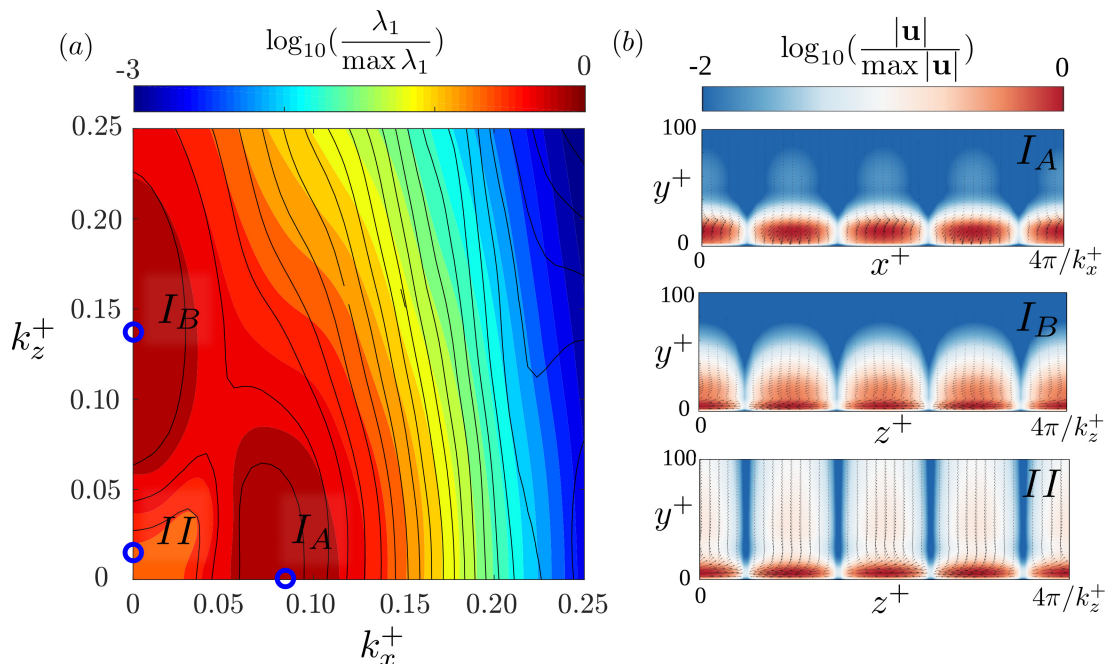
FIG. 9. (a) The largest eigenvalue of the Hessian matrix $\hat{\mathcal{H}}(k_x, k_z; t_m)$ for observing $\partial u/\partial y|_{\text{wall}}$ instantaneously at $t_m^+ = 20$. Color and line contours are for $Re_\tau = 180$ and 590, respectively. The eigenvalues are normalized by the supremum. (b) Visualizations of the leading eigenfunctions of the Hessian at selected wavenumber pairs: $(k_x^+, k_z^+) = (0.08, 0)$, $(0, 0.14)$, $(0, 0.02)$ for modes I_A, I_B and II. Contours show the velocity magnitude and arrows are in-plane velocity components.

other words, these flow structures will be the easiest to reconstruct from observations. The associated eigenvalues quantify the curvature of the cost function along the corresponding eigen-direction. (ii) Eigen-analysis of the Hessian matrix is equivalent to a proper-orthogonal decomposition (POD) of the adjoint field $\boldsymbol{u}^\dagger$. As such, the leading eigenfunctions are the dominant shapes of the domain of dependence. In the forward sense, these eigenfunctions are also the optimal perturbations that will lead to the most significant change in wall signals at t_m, and the corresponding sensitivity is quantified by the eigenvalues.

A sample eigenspectrum and eigenfunctions of the Hessian for observing $\partial u/\partial y|_{\text{wall}}$ is shown in figure 9. Exploiting periodicity in the horizontal x-z plane, the Hessian is expressed in Fourier space,

$$\hat{\mathcal{H}}(k_x, k_z; t_m) = \frac{1}{S}\int_S \hat{\boldsymbol{u}}^\dagger(k_x, k_z, \tau = t_m; \boldsymbol{x}_m, t_m)\hat{\boldsymbol{u}}^\dagger(k_x, k_z, \tau = t_m; \boldsymbol{x}_m, t_m)dx_m dz_m. \tag{22}$$

The largest eigenvalue of $\hat{\mathcal{H}}(k_x, k_z; t_m)$ at each wavenumber pair is plotted in figure 9a. Two Reynolds numbers $Re_\tau = 180$ and 590 are considered, and their eigenspectra (color and line contours) coincide when the wavenumber are scaled in viscous units. The maximum eigenvalue decays with increasing magnitude of the horizontal wavenumber vector (k_x^+, k_z^+), and the separation of eigenvalues over three orders of magnitudes demonstrates the ill-conditioned nature of the Hessian. Two local peaks marked I_A and I_B correspond to spanwise and streamwise rolls. Another important point is marked II, which represents the energetic large-scale structures with $\lambda_z^+ = \mathcal{O}(300)$. The mode shapes of I_A, I_B and II are plotted in figure 9b.

Within the considered time horizon, $t_m^+ = 20$, the initial disturbances that can most effectively influence wall measurements have large wavenumbers, and the corresponding eigenvectors are clustered near the wall (modes I_A and I_B in figure 9b). These modes are not energetic in channel-flow turbulence, but are nonetheless the directions that adjoint-variational data assimilation will attempt to reconstruct first. Another important remark is that, beyond the buffer layer, most eigenfunctions are vanishing [see 45, for more details], which implies that wall observations are not very sensitive to the turbulence far from the wall. A very important exception is in a small region near mode II, where the eigenfunctions remain finite in the core of the channel. These modes represent the sensitivity of wall signals to outer large-scale structures. This sensitivity arises because these streamwise-elongated motions are immune to shear sheltering [46, 47], have a footprint in the near-wall region and modulate the near-wall structures [48–51]. As a result, they have a signature encoded in the wall stress. The analysis of the Hessian eigenspectrum thus provides a rigorous explanation of the empirical findings by earlier studies [15, 19, 43]: while wall observations can be decoded to discover all the turbulent scales near the wall, only the outer large-scale motions can be accurately estimated from

wall stresses.

The present discussion has focused on an instantaneous observation, at $t_m^+ = 20$. In practice, observations are accumulated over a time horizon that may extend over a longer duration. Due to the Lyapunov behaviour of the *adjoint* system, the adjoint field amplifies exponentially in reverse time at the same rate as the forward system. The notion of the adjoint chaos is, however, different from the familiar forward problem: To start, the initial guess of the flow state in a data assimilation procedure is generally far from the truth, and the associated gradient of the cost function is determined by the adjoint system. The exponential amplification of the adjoint in backward time can lead to a very large gradient, which imposes a severe restriction on the step size in gradient descent optimization. Even if the initial estimate is near optimality, the Hessian matrix, which is represented by the adjoint field starting from an observation kernel, has exponentially amplifying eigenvalues with observation time. This property translates into a cost function landscape with extreme curvatures, which can thwart the search for the global optimum. More subtle but very important is an increase in the separation between the largest and smallest eigenvalues of the Hessian, and hence an increase of its condition number. As a result, solving the inverse problem becomes progressively more ill conditioned and errors in observations can strongly obscure the reconstruction of the initial state. Effective strategies to address these difficulties include limiting the assimilation horizon based on knowledge of the Lyapunov timescale [15, 52], and for longer horizons adopting a sliding window [20] or a cycling scheme [29]. These techniques should also be combined with strategies for optimizing sensor placement and weighting [17, 18] in order to maximize the accuracy of predictions.

IV. CONCLUSION

Data assimilation combines experimental observations with numerical simulations in order to achieve high-fidelity predictions of the entire flow state. The data-infused simulations improve the accuracy of the computations and reduce the uncertainty of predictions. At the same time, experimental measurements are augmented with non-intrusive access to a much higher resolution than the sensor data. These benefits have been demonstrated by applying adjoint-variational data assimilation in order to reconstruct a nonlinearly saturated flow and wall turbulence, from limited observations. When the flow configuration admits multiple states, such as in circular Couette flow, the observed one may originate from a complicated history and particulars of the experimental setup. Data assimilation can successfully and accurately reconstruct the true state from sparse observations, such as wall stress measurements, without the requirement of matching the detailed experimental setup and protocol [26]. In wall turbulence, starting from sparse velocity data, adjoint-variational assimilation provided reconstructions of the full velocity and vorticity fields that are nearly perfectly correlated with the truth and robust to observation noise [15].

The importance of data assimilation extends beyond augmenting observations. In particular, adjoint-variational methods provide a unique perspective on flow physics that arises from the process of decoding observations. The adjoint field quantifies the sensitivity of the observation kernel to changes in the flow state, and hence the domain of dependence of the sensor. This view motivates a shift in focus from clustering measurements in order to locally observe all the scales to optimally distributing observations in order to ensure sensitivity to the entire state. In order to guarantee accurate reconstruction of the flow, a critical data resolution must be satisfied, whereby the domains of dependence of adjacent observation sites overlap. This resolution was shown to be on the order of the Taylor microscale of turbulence.

The search for the true flow is performed by minimizing a cost function that is defined in terms of the disparity between model predictions and observations. Success must therefore contend with the landscape of this cost function, which is mountainous for nonlinear chaotic systems. We have shown an example of the cost landscape explicitly using the Lorenz system—a practically impossible task in high dimensions such as for turbulence. Luckily the adjoint system provides a solution [45]: At optimality, the Hessian matrix of the cost can be computed relatively efficiently from the adjoint field, and its eigen-decomposition provides important information. The leading eigenfunctions are the flow structures that have the most impact on observations. The eigenvalues are the curvatures of the cost landscape along the associated eigen-directions, and describe the difficulty of approaching optimality along those directions. As such, from the perspective of data assimilation, or the inverse problem, the leading eigenfunction are the most accurate to decode from observations.

The above ideas were applied to the interpretation of observations of the wall stress in turbulent channel flow. Within a short time horizon, these observations are most sensitive to near-wall small-scale perturbations, and strongly affected by outer large-scale structures. At longer observation time horizons, the sensitivity amplifies exponentially due to the chaotic nature of the adjoint system, leading to a more mountainous cost function, larger gradient, and a deteriorating ill-conditioned Hessian. Therefore, the inverse problem becomes progressively more difficult to solve, and errors in observations can strongly contaminate the estimated initial state. These difficulties restrict assimilation time horizon to the order of the Lyapunov timescale.

The difficulty of turbulence reconstruction can be alleviated by designing optimal measurement layouts, which can be guided by the Hessian analysis. The weighting of observations and spatio-temporal placement of sensors can be optimized to mitigate undesirable features of the cost landscape. For example, in transitional or turbulent flows earlier observations can be favorably weighted in order to reduce extreme curvature, and thus improve the estimation accuracy of the initial state. In turbulent flows, the sensor placement can be designed to mitigate the poor condition of the Hessian [18]. These strategies have been applied to prediction of high-speed boundary layer transition and scalar source reconstruction in turbulent environments [16, 17].

At much higher Reynolds numbers, the challenges examined herein become more severe and new ones arise. The higher Lyapunov exponent will further restrict the duration of the assimilation window. In addition, a larger number of observations, or sensors, will be required at higher Reynolds numbers, which can render estimation of the entire flow field prohibitive both experimentally and computationally. Measurements can therefore be concentrated in a small region of interest, and the data assimilation can reconstruct the flow within this sub-domain along with the effective boundary conditions due to the influence of the larger configuration. Finally, turbulence modeling may be required in the forward and adjoint equations, which introduces uncertainty with respect to model parameters. All these considerations are exciting future avenues of inquiry.

Furthermore, the results from the adjoint-variational approach provide a benchmark for assessing emerging data-assimilation techniques, for example recent machine learning strategies that integrate data and physics-based, or equations, constraints. Notable among the growing suite of these methods are physics-informed neural networks [21], evolutionary deep neural networks [EDNN which is pronounced Eden, see 22], and Deep Operator Networks [DeepONet, see 53]. A systematic comparison of adjoint-variational and machine-learning data assimilation techniques is currently under investigation, and future efforts may combine the two classes of methods in multi-fidelity data-assimilation strategies.

Declaration of interests. The authors report no conflict of interest.

[1] P. K. Yeung, K. R. Sreenivasan, and S. B. Pope, Effects of finite spatial and temporal resolution in direct numerical simulations of incompressible isotropic turbulence, Phys. Rev. Fluids **3**, 064603 (2018).

[2] M. Lee and R. D. Moser, Direct numerical simulation of turbulent channel flow up to $Re_\tau \approx 5200$, J. Fluid Mech. **774**, 395 (2015).

[3] O. Marxen, G. Iaccarino, and T. E. Magin, Direct numerical simulations of hypersonic boundary-layer transition with finite-rate chemistry, J. Fluid Mech. **755**, 35 (2014).

[4] A. Esteghamatian and T. A. Zaki, Viscoelasticity and the dynamics of concentrated particle suspension in channel flow, J. Fluid Mech. **901**, A25 (2020).

[5] S. J. Lee and T. A. Zaki, Simulations of natural transition in viscoelastic channel flow, J. Fluid Mech. **820**, 232-262 (2017).

[6] J. H. Seo, V. Vedula, T. Abraham, and R. Mittal, Multiphysics computational models for cardiac flow and virtual cardiography, Int. J. Num. Meth. Biomed. Eng. **29**, 850 (2013).

[7] L. Nicolaou and T. Zaki, Direct numerical simulations of flow in realistic mouth-throat geometries, J. Aerosol Sci. **57**, 71 (2013).

[8] X. I. Yang and C. Meneveau, Recycling inflow method for simulations of spatially evolving turbulent boundary layers over rough surfaces, J. Turbul. **17**, 75 (2016).

[9] X. Wu, Inflow turbulence generation methods, Annu. Rev. Fluid Mech. **49**, 23 (2017).

[10] S. G. Huisman, R. C. Van Der Veen, C. Sun, and D. Lohse, Multiple states in highly turbulent taylor–couette flow, Nat. Commun. **5**, 1 (2014).

[11] A. J. Smits, B. J. McKeon, and I. Marusic, High-Reynolds number wall turbulence, Annu. Rev. Fluid Mech. **43**, 353 (2011).

[12] J. Wang, C. Zhang, and J. Katz, Gpu-based, parallel-line, omni-directional integration of measured pressure gradient field to obtain the 3D pressure distribution, Exp. Fluids **60**, 1 (2019).

[13] A. Stuart and K. Zygalakis, *Data Assimilation: A Mathematical Introduction*, Tech. Rep. (Oak Ridge National Lab.(ORNL), Oak Ridge, TN (United States), 2015).

[14] V. Mons, J.-C. Chassaing, T. Gomez, and P. Sagaut, Reconstruction of unsteady viscous flows using data assimilation schemes, J. Comput. Phys. **316**, 255 (2016).

[15] M. Wang and T. A. Zaki, State estimation in turbulent channel flow from limited observations, J. Fluid Mech. **917**, A9 (2021).

[16] 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).

[17] V. Mons, Q. Wang, and T. A. Zaki, Kriging-enhanced ensemble variational data assimilation for scalar-source identification in turbulent environments, J. Comput. Phys. **398**, 108856 (2019).

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

[19] M. P. Encinar and J. Jiménez, Logarithmic-layer turbulence: A view from the wall, Phys. Rev. Fluids **4**, 114603 (2019).

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

[21] M. Raissi, P. Perdikaris, and G. E. Karniadakis, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, J. Comput. Phys. **378**, 686 (2019).

[22] Y. Du and T. Zaki, Evolutional deep neural network, arXiv preprint arXiv:2103.09959 (2021).

[23] G. Evensen, Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics, J. Geophys. Res. Oceans **99**, 10143 (1994).

[24] T. Suzuki, Reduced-order Kalman-filtered hybrid simulation combining particle tracking velocimetry and direct numerical simulation, J. Fluid Mech. **709**, 249-288 (2012).

[25] F.-X. L. Dimet and O. Talagrand, Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects, Tellus A: Dyn. Meteorol. Oceanogr. **38**, 97 (1986).

[26] M. Wang, Q. Wang, and T. A. Zaki, Discrete adjoint of fractional-step incompressible Navier-Stokes solver in curvilinear coordinates and application to data assimilation, J. Comput. Phys. https://doi.org/10.1016/j.jcp.2019.06.065 (2019).

[27] N. Chandramoorthy, P. Fernandez, C. Talnikar, and Q. Wang, Feasibility analysis of ensemble sensitivity computation in turbulent flows, AIAA J. **57**, 4514 (2019).

[28] G. Eyink, T. Haine, and D. Lea, Ruelle's linear response formula, ensemble adjoint schemes and Lévy flights, Nonlinearity **17**, 1867 (2004).

[29] M. Fisher and H. Auvinen, Long window 4d-var, in Seminar on Data assimilation for atmosphere and ocean, 6-9 September 2011, ECMWF (ECMWF, Shinfield Park, Reading, 2012) pp. 189–202.

[30] J. Nocedal, Updating quasi-Newton matrices with limited storage, Math. Comput. **35**, 773 (1980).

[31] T. A. Zaki, From streaks to spots and on to turbulence: exploring the dynamics of boundary layer transition, Flow Turbul. Combust. **91**, 451 (2013).

[32] O. Marxen and T. A. Zaki, Turbulence in intermittent transitional boundary layers and in turbulence spots, J. Fluid Mech. **860**, 350 (2019).

[33] Z. Wu, T. A. Zaki, and C. Meneveau, High-Reynolds-number fractal signature of nascent turbulence during transition, Proc. Natl. Acad. Sci. U.S.A. **117**, 3461 (2020).

[34] O. Reynolds, XXIX. An experimental investigation of the circumstances which determine whether the motion of water shall be direct or sinuous, and of the law of resistance in parallel channels, Phil. Trans. R. Soc. , 935 (1883).

[35] K. Avila, D. Moxey, A. de Lozar, M. Avila, D. Barkley, and B. Hof, The onset of turbulence in pipe flow, Science **333**, 192 (2011).

[36] E. N. Lorenz, Deterministic nonperiodic flow, J. Atmos. Sci. **20**, 130 (1963).

[37] B. Saltzman, Finite amplitude free convection as an initial value problem-I, J. Atmos. Sci. **19**, 329 (1962).

[38] N. Nikitin, Characteristics of the leading Lyapunov vector in a turbulent channel flow, J. Fluid Mech. **849**, 942 (2018).

[39] S. Abrahamson and S. Lonnes, Uncertainty in calculating vorticity from 2D velocity fields using circulation and least-squares approaches, Exp. Fluids **20**, 10 (1995).

[40] G. L. Eyink, A. Gupta, and T. A. Zaki, Stochastic Lagrangian dynamics of vorticity. Part 1. General theory for viscous, incompressible fluids, J. Fluid Mech. **901**, A2 (2020).

[41] G. L. Eyink, A. Gupta, and T. A. Zaki, Stochastic Lagrangian dynamics of vorticity. Part 2. Application to near-wall channel-flow turbulence, J. Fluid Mech. **901**, A3 (2020).

[42] T. R. Bewley and B. Protas, Skin friction and pressure: the footprints of turbulence, Phys. D: Nonlinear Phenom. **196**, 28 (2004).

[43] T. Suzuki and Y. Hasegawa, Estimation of turbulent channel flow at $Re_\tau = 100$ based on the wall measurement using a simple sequential approach, J. Fluid Mech. **830**, 760 (2017).

[44] C. W. Rowley, Model reduction for fluids, using balanced proper orthogonal decomposition, Int. J. Bifurcat. Chaos **15**, 997 (2005).

[45] Q. Wang, M. Wang, and T. Zaki, From wall observations to turbulence: The difficulty of flow reconstruction, arXiv preprint arXiv:2106.09169 (2021).

[46] J. Hunt and P. Durbin, Perturbed vortical layers and shear sheltering, Fluid Dyn. Res. **24**, 375 (1999).

[47] T. A. Zaki and S. Saha, On shear sheltering and the structure of vortical modes in single-and two-fluid boundary layers, J. Fluid Mech. **626**, 111 (2009).

[48] R. Mathis, N. Hutchins, and I. Marusic, Large-scale amplitude modulation of the small-scale structures in turbulent boundary layers, J. Fluid Mech. (2009).

[49] H. Abe, H. Kawamura, and H. Choi, Very large-scale structures and their effects on the wall shear-stress fluctuations in a turbulent channel flow up to re τ= 640, J. Fluids Eng. **126**, 835 (2004).

[50] J. Hwang, J. Lee, H. J. Sung, and T. A. Zaki, Inner-outer interactions of large-scale structures in turbulent channel flow, J. Fluid Mech. **790**, 128 (2016).

[51] J. You and T. A. Zaki, Conditional statistics and flow structures in turbulent boundary layers buffeted by free-stream

disturbances, J. Fluid Mech. **866**, 526-566 (2019).

[52] 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** (2020).

[53] P. C. Di Leoni, L. Lu, C. Meneveau, G. Karniadakis, and T. A. Zaki, Deeponet prediction of linear instability waves in high-speed boundary layers, arXiv preprint arXiv:2105.08697 (2021).