Reduced-order description of transient instabilities and computation of finite-time Lyapunov exponents

Hessam Babaee,[1,2] Mohamad Farazmand,[1] George Haller,[3] and Themistoklis P. Sapsis[1,a]

[1]*Department of Mechanical Engineering, MIT Cambridge, Cambridge, Massachusetts 02139, USA*
[2]*Department of Mechanical Engineering and Materials Science, University of Pittsburgh, Pittsburgh, Pennsylvania 15261, USA*
[3]*Department of Mechanical and Process Engineering, ETH Zurich, 8092 Zurich, Switzerland*

High-dimensional chaotic dynamical systems can exhibit strongly transient features. These are often associated with instabilities that have a finite-time duration. Because of the finite-time character of these transient events, their detection through infinite-time methods, e.g., long term averages, Lyapunov exponents or information about the statistical steady-state, is not possible. Here, we utilize a recently developed framework, the Optimally Time-Dependent (OTD) modes, to extract a time-dependent subspace that spans the modes associated with transient features associated with finite-time instabilities. As the main result, we prove that the OTD modes, under appropriate conditions, converge exponentially fast to the eigendirections of the Cauchy–Green tensor associated with the most intense finite-time instabilities. Based on this observation, we develop a reduced-order method for the computation of finite-time Lyapunov exponents (FTLE) and vectors. In high-dimensional systems, the computational cost of the reduced-order method is orders of magnitude lower than the full FTLE computation. We demonstrate the validity of the theoretical findings on two numerical examples. *Published by AIP Publishing.* [http://dx.doi.org/10.1063/1.4984627]

For a plethora of dynamical systems transient phenomena—usually associated with finite-time instabilities—define the most important aspects of the response. Examples include rare chaotic bursts in turbulent systems, extreme events in nonlinear waves, and strong transients in complex networks. In such systems, the rare transient responses coexist with a high dimensional chaotic attractor where the system spends most of its lifetime. Because of the rare character of the transient dynamics, it is particularly challenging to extract modes that capture them using traditional methods for order reduction and base selection. Here, we prove that a recently developed method, the optimally time-dependent (OTD) modes, provides a time-dependent basis that spans the directions associated with the most intense instabilities, even if these have short-time character. Therefore, the derived modes encode the strongly time-dependent features of the associated instabilities. We demonstrate the developed framework through the computation of finite-time Lyapunov exponents (FTLE) utilizing only the projected dynamics along the derived modes.

I. INTRODUCTION

For a plethora of dynamical systems transient phenomena—usually associated with finite-time instabilities—define the most important aspects of the response. Examples include chaotic fluid systems,[4,16,17] nonlinear waves,[8,9] and networks.[7,38] For the analysis of such systems and the formulation of prediction and control algorithms, it is critical to identify modes associated with these strongly transient features. A first important challenge towards identifying these modes is the high-dimensionality of the response. This high-dimensionality does not necessarily imply that the transient features are also high-dimensional. In fact, it is often the case that the modes associated with transient features are very few. However, because of the broad spectrum of the response it is hard to identify them using energy-based criteria. An additional challenge is associated with the intrinsically time-dependent character of these features which is, in general, non-periodic and makes it essential to consider arbitrary time-dependence for these modes as well.

In the context of dynamical systems, numerous approaches have been developed to characterize the transient features associated with non-periodic behavior. The notion of Lagrangian Coherent Structures (LCS) emerged from the fluid dynamics problem of mixing as a general method to characterize dynamical systems with arbitrary time-dependence.[20,22,26] The method quantifies finite-time instabilities through the explicit computation of the maximum eigenvalue of the Cauchy–Green tensor on every location of the phase space. For low-dimensional systems, it provides a complete and unambiguous description of the finite-time instabilities by identifying parts of the phase space associated with such responses. However, given the fast growth of the computational cost with respect to the phase space dimensionality it is not possible to apply it for very high- or infinite-dimensional systems in order to characterize finite-time instabilities.

For such problems, other approaches focusing more on the description of the system attractor rather than the whole phase space have also been developed in the context of

a)Author to whom correspondence should be addressed: sapsis@mit.edu. Tel.: (617) 324-7508. Fax: (617) 253-8689.

uncertainty quantification and model order reduction. Specifically, the notion of dynamical orthogonality[5,6,33,34] allows for the computation of time-dependent, dynamically orthogonal (DO) modes that lead to reduced-order description of chaotic attractors with low-intrinsic dimensionality but with strong time-dependence.[31,32] The same notion can be combined with stochastic closures[35,37] in order to quantify statistics in time-dependent subspaces for turbulent dynamics systems with very broad spectra (or high intrinsic dimensionality) and strongly transient features.[27,36] In all of these studies, the effectiveness of the derived DO modes in capturing transient responses was demonstrated numerically and very few theoretical studies related to this problem were constrained in very specific setups. One such case is the study of linear parabolic (stable) equations in Ref. 28 where it was shown that the DO modes tend to capture the most energetic directions. However, there is no general result indicating where the DO modes converge for an arbitrary (nonlinear) dynamical system.

To circumvent this limitation, a minimization principle was introduced in Ref. 2 that allows for the derivation of equations that evolve a time-dependent set of modes, the Optimally Time Dependent (OTD) modes, along a given trajectory of the system. These modes are identical to those obtained from the DO equations in the deterministic limit, i.e., for the case where the stochastic energy goes to zero. To this end, the OTD modes can be seen as the deterministic analog of DO modes. For sufficiently long times where the system reaches an equilibrium, it was proven in Ref. 2 that the OTD modes converge to the most unstable directions of the system in the asymptotic limit. Moreover, the same modes were utilized in Ref. 16 to formulate predictors of extreme events for Navier-Stokes equations and nonlinear wave problems.

Despite their success in numerically capturing transient phenomena associated with finite-time instabilities, the exact relation between OTD modes (or DO modes) and finite-time instabilities remains an open problem. In this work, we provide a rigorous link between the OTD modes and finite-time instabilities, by showing that under general conditions the OTD modes converge exponentially fast to the eigendirections of the Cauchy–Green tensor associated with the largest eigenvalues, i.e., with the largest finite-time Lyapunov exponents. Thus, we prove that the OTD modes are able to extract over a finite-time interval the modes associated with the most intense finite-time instabilities or the most pronounced transient phenomena. Note, that such convergence does not depend on the dimensionality of the system or the intrinsic dimensionality of the attractor.

Apart from the fundamental implications of this result on the formulation of reduced-order algorithms for the prediction and control of transient phenomena, there is a direct application on the computation of the maximum finite-time Lyapunov exponents for high- or infinite-dimensional systems. Specifically, using the OTD modes we formulate a reduced-order framework that allows, under mild conditions, the computation of finite-time Lyapunov Exponents for general dynamical systems without having to compute the full Cauchy–Green tensor. We demonstrate the derived algorithm

in two systems, the Arnold-Beltrami-Childress (ABC) flow and the Six-dimensional Charney-DeVore (CDV) model with regime transitions. In both cases, we examine the effectiveness of the reduced-order method and study limitations and applicability.

II. PRELIMINARIES AND NOTATION

We consider the system of differential equations

$$\dot{z} = f(z,t), \quad z \in \mathbb{R}^n, \quad t \in I = [t_0, t_0 + T], \quad (1)$$

where $f : U \times I \to \mathbb{R}^n$ is a sufficiently smooth vector field. Let $F_{t_0}^t$ with $t \in I$ denote the associated flow map,

$$F_{t_0}^t : U \to U$$
$$z_0 \mapsto z(t; t_0, z_0), \quad (2)$$

where $z(t; t_0, z_0)$ is a trajectory of system (1) with the initial condition z_0. The linearized system around the trajectory $z(t) \equiv z(t; t_0, z_0)$ satisfies the equation of variations,

$$\dot{v} = L(z(t), t)v, \quad v(t) \in \mathbb{R}^n, \quad t \in I = [t_0, t_0 + T], \quad (3)$$

where $L(z,t) := \nabla_z f(z,t)$. The deformation gradient $\nabla F_{t_0}^t$ is the fundamental solution matrix for the linearized dynamics such that the solutions $v(t) \equiv v(t; t_0, v_0)$ of the equation of variations satisfy

$$v(t, t_0, z_0) = \nabla F_{t_0}^t(z_0)v_0, \quad \text{with } t \in I = [t_0, t_0 + T]. \quad (4)$$

A. Finite-time Lyapunov exponents

To measure the growth of infinitesimal perturbations in the phase space, we use the *right Cauchy–Green strain tensor*

$$C_{t_0}^t = (\nabla F_{t_0}^t)^T \nabla F_{t_0}^t,$$

and the *left Cauchy–Green strain tensor*

$$B_{t_0}^t = \nabla F_{t_0}^t (\nabla F_{t_0}^t)^T, \quad (5)$$

where T denotes matrix transposition. Let $\xi(t, t_0, z_0) \in \mathbb{R}^n$ denote the eigenvectors of the right Cauchy–Green strain tensor, so that

$$C_{t_0}^t(z_0)\xi_i(t, t_0, z_0) = \lambda_i(t, t_0, z_0)\xi_i(t, t_0, z_0), \quad i = 1, \ldots, n,$$

where $\lambda_i(t, t_0, z_0)$ are the corresponding eigenvalues. Similarly, denote the eigenvectors of the left Cauchy–Green strain tensor by $\eta_i(t, t_0, z_0) \in \mathbb{R}^n$ and their corresponding eigenvalues by $\mu_i(t, t_0, z_0)$, so that

$$B_{t_0}^t(z_0)\eta_i(t, t_0, z_0) = \mu_i(t, t_0, z_0)\eta_i(t, t_0, z_0), \quad i = 1, \ldots, n.$$

When no confusion may arise, we omit the dependence of the eigenvalues and eigenvectors on (t, t_0, z_0) for notational simplicity. Since the Cauchy–Green strain tensors are symmetric and positive definite, their eigenvalues are real and

positive, $\lambda_i, \mu_i \in \mathbb{R}^+$. Furthermore, the eigenvectors are orthogonal, i.e.,

$$\langle \xi_i, \xi_j \rangle = \langle \eta_i, \eta_j \rangle = \delta_{ij}, \quad i, j = 1, \ldots, n,$$

where $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product.

It is straightforward to show that the right and left tensors have the same eigenvalues, $\lambda_i = \mu_i$. Specifically, considering the definition of the right Cauchy–Green tensor eigenvalues, $(\nabla F_{t_0}^t)^T \nabla F_{t_0}^t \xi_i = \lambda_i \xi_i$, we multiply the equation with $\nabla F_{t_0}^t$ from the left to obtain

$$B_{t_0}^t (\nabla F_{t_0}^t \xi_i) = \lambda_i (\nabla F_{t_0}^t \xi_i). \qquad (6)$$

Thus, the left and right Cauchy–Green strain tensors have the same set of eigenvalues. Finally, using the singular value decomposition of the deformation gradient, one can show the well-known relation that

$$\nabla F_{t_0}^t \xi_i = \sqrt{\lambda_i} \eta_i, \quad i = 1, \ldots, n, \qquad (7)$$

(see, e.g., Ref. 23). In the following, we order the Cauchy–Green eigenvalues in a descending order,

$$\lambda_1 > \lambda_2 > \cdots > \lambda_n > 0. \qquad (8)$$

The finite-time Lyapunov exponents (FTLEs) of system (1) corresponding to the trajectory $z(t; t_0, z_0)$ and the time interval $[t_0, t_0 + T]$ are defined as

$$\Lambda_i(t_0 + T, t_0, z_0) = \frac{1}{T} \log \sqrt{\lambda_i(t_0 + T, t_0, z_0)},$$
$$i = 1, 2, \ldots, n. \qquad (9)$$

Note that FTLEs are well-defined for any finite integration time T. Under certain assumptions, the limit $T \to \infty$ also exists.[30]

B. Optimally time-dependent (OTD) modes

We give an overview of the OTD modes[2] for general dynamical systems. The OTD modes represent a reduced-order set of r time-dependent orthonormal modes, $U(t) = [u_1(t), u_2(t), \ldots, u_r(t)]$, that minimize the difference between the action of the infinitesimal propagator $\nabla F_t^{t+\delta t}$ on $u_i(t)$ and its value at time $t + \delta t$. More specifically, we seek to minimize the functional

$$\mathcal{F} = \frac{1}{(\delta t)^2} \sum_{i=1}^{r} ||u_i(t + \delta t) - \nabla F_t^{t+\delta t} u_i(t)||^2, \quad \delta t \to 0, \quad (10)$$

with the constraint that the time-dependent basis satisfies the orthonormality conditions

$$\langle u_i(t), u_j(t) \rangle = \delta_{ij}, \quad i, j = 1, \ldots, r. \qquad (11)$$

Using Taylor series expansions, we have

$$u_i(t + \delta t) = u_i(t) + \delta t \, \dot{u}_i + \mathcal{O}(\delta t^2), \qquad (12)$$

$$\nabla F_t^{t+\delta t} = I + \delta t \, L(z(t), t) + \mathcal{O}(\delta t^2), \qquad (13)$$

where I is the identity matrix. Replacing the above equations into the functional (10) results in

$$\mathcal{F}(\dot{u}_1, \dot{u}_2, \ldots, \dot{u}_r) = \sum_{i=1}^{r} \left\| \frac{\partial u_i(t)}{\partial t} - L(z(t), t) u_i(t) \right\|^2. \qquad (14)$$

We emphasize that the minimization of the function (14) is considered only with respect to the time-derivative (rate of change) of the basis, $\dot{U}(t)$, instead of the basis $U(t)$ itself. This is because we do *not* want to optimize the subspace that the operator is acting on, but rather find an optimal set of vectors, $\dot{U}(t)$, that best approximates the linearized dynamics in the subspace U. We then solve the resulted equations and compute $U(t)$. We will refer to these modes as the *optimally time-dependent* (OTD) modes, and the space that these modes span as the OTD subspace. By utilizing the minimization principle and taking into account the orthonormality constraint, we obtain the following theorem (proved in Ref. 2.)

Theorem 2.1. *A one parameter family of vectors $u_i(t) \in \mathbb{R}^n$, $i = 1, 2, \ldots, r$, minimizes the functional (14) and satisfies the orthonormality condition (11) if and only if*

$$\frac{\partial U}{\partial t} = LU - UU^T LU, \qquad (15)$$

where $U(t) = [u_1(t), u_2(t), \ldots, u_r(t)] \in \mathbb{R}^{n \times r}$.

Equation (15) is the evolution equation for the OTD modes. It was proven in Ref. 2 that for the case of a time independent operator, L, the subspace spanned by the columns of $U(t)$ converges asymptotically to the modes associated with the most unstable directions of the operator L.

The OTD modes were also used to capture the transient non-normal growth of instabilities. In Fig. 1, we recall a system from Ref. 2 exhibiting a non-normal growth (cf. Ref. 2 for the details). A typical trajectory of the system, shown in the left panel, has an almost periodic behavior where each cycle exhibits three distinct regimes: (A) an exponential growth in the z_3 direction, (B) a non-normal growth in the $z_1 - z_2$ plane with simultaneous decay in the z_3 direction, and (C) exponential decay in the $z_1 - z_2$ plane to the origin. This configuration allows for the repeated occurrence of exponential (regime A) and non-normal (regime B) instabilities.

Due to the exponential instability close to the origin, the system undergoes chaotic transitions between positive and negative values of z_3. We compute the instantaneous growth rate corresponding to the direction of a single OTD mode. The OTD mode initially captures the severe exponential growth and subsequently captures the non-normal growth. On the other hand, the real part of the eigenvalues of the full linearized operator can only capture the exponential growth, even in regimes where it is not relevant, while they completely miss the non-normal growth. The maximum largest singular value σ of the matrix L exhibits a similar behavior.

The OTD modes share some fundamental characteristics of other stability measures. In Sec. III, we show the relation between the OTD modes and the finite-time Lyapunov vectors. In Appendix A, we also discuss the connection between them and the Dynamically Orthogonal (DO) modes
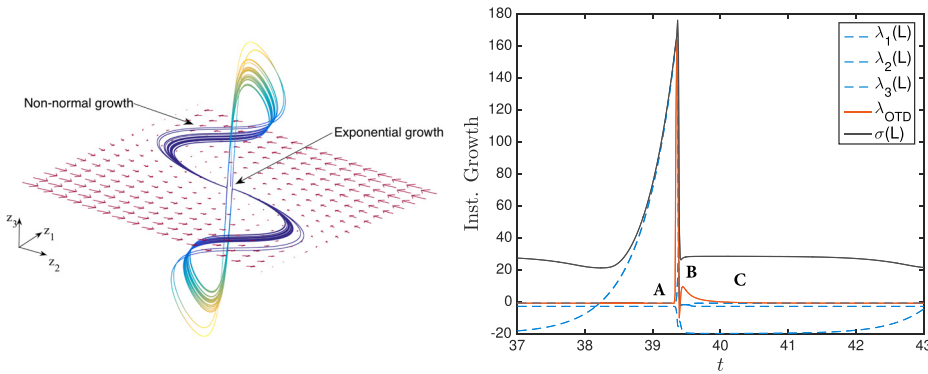
FIG. 1. An example of the OTD modes on the analysis of transient instabilities. Left: A trajectory of the dynamical system colored according to the state variable z_3. The non-normal vector field for $z_3 = 0$ is also shown. Right: The three eigenvalues of the linearized operator are plotted with blue dashed curves, while the growth rate of the single OTD mode is shown with orange color. The maximum eigenvalue of the singular value decomposition of the linearization along the trajectory is also shown.

introduced in Ref. 33. Their connection with covariant Lyapunov vectors[18] remains to be explored.

III. ALIGNMENT OF THE OTD MODES WITH THE DOMINANT CAUCHY–GREEN STRAIN EIGENVECTORS

Here, we prove that under proper assumptions, the OTD modes not only capture the unstable directions of steady linear operators, but also the transient instabilities associated with time-dependent operators. More specifically, we prove that the OTD subspace converges exponentially fast to the eigenspace associated with the dominant eigenvalues of the Cauchy–Green strain tensor. The latter describes the finite-time growth of infinitesimal perturbations in the phase space. For what follows, we will need to measure the distance between two subspaces of the same dimensionality. To this end, we need the following definitions.

Definition 3.1. *Two r-dimensional linear subspaces of $\mathbb{R}^n$ spanned by the columns of $U \in \mathbb{R}^{n \times r}$ and $V \in \mathbb{R}^{n \times r}$ are equivalent if there is an invertible matrix $R \in \mathbb{R}^{r \times r}$ such that $U = VR$.*

The next definition provides a quantity for measuring the "angle" between two subspaces.

Definition 3.2. *For two r-dimensional linear subspaces of $\mathbb{R}^n$ spanned by the columns of $U = [u_1, u_2, ..., u_r] \in \mathbb{R}^{n \times r}$ and $V = [v_1, v_2, ..., v_r] \in \mathbb{R}^{n \times r}$, we define the distance function*

$$\gamma_{U,V} = \frac{||U^T V||}{r^{1/2}}, \qquad (16)$$

where $|| \cdot ||$ denotes the Frobenius norm.

Note that the entries of $U^T V$ are the inner products between u_i's and v_j's, i.e., $(U^T V)_{ij} = \langle u_i, v_j \rangle$. The following lemma shows that the equivalence of two subspaces can be deduced from the distance $\gamma_{U,V}$.

Lemma 3.1. *Given a subspace defined by the orthonormal columns of $U \in \mathbb{R}^{n \times r}$, and a subspace defined by the unit-length (but not necessarily orthonormal) columns of $V \in \mathbb{R}^{n \times r}$, then we always have $\gamma_{U,V} \leq 1$. Moreover, the two subspaces spanned by the columns of U and V are equivalent if and only if $\gamma_{U,V} = 1$.*

Proof. By the orthogonal projection theorem, there are $h_i \in \mathbb{R}^r$ and $b_i \in \mathbb{R}^n$ such that $v_i = U h_i + b_i$ and $U^T b_i = 0$ for $i = 1, 2, ..., r$. Note that since v_i are unit length, we have

$$1 = ||v_i||^2 = ||h_i||^2 + ||b_i||^2, \quad i = 1, 2, ..., r.$$

Defining $H = [h_1 | \cdots | h_r]$ and $B = [b_1 | \cdots | b_r]$, we have $V = UH + B$. This implies $U^T V = H$ and hence

$$||U^T V||^2 = ||H||^2 = \sum_{i=1}^{r} ||h_i||^2 = r - \sum_{i=1}^{r} ||b_i||^2 = r - ||B||^2.$$

Therefore, $\gamma_{U,V}^2 = 1 - ||B||^2 / r$. Note that the subspaces $\text{col}(U)$ and $\text{col}(V)$ are equal if and only if $||B|| = 0$. Therefore, $\text{col}(U) = \text{col}(V)$ if and only if $\gamma_{U,V} = 1$. We also note that if the two subspaces are not equivalent $\gamma_{U,V} < 1$. $\square$

To prove the main theorem, it is more convenient to work with the subspace spanned by the linearized flow (3), instead of the OTD subspace. The following result states that the two subspaces are equivalent.

Theorem 3.1. *Let $V(t) \in \mathbb{R}^{n \times r}$ solve the equation of variations (3) and $U(t) \in \mathbb{R}^{n \times r}$ with $U^T U = I$ solve the OTD equation (15). Assume that the two subspaces spanned by the columns of U and V are initially equivalent, i.e., there is an invertible matrix $T_0 \in \mathbb{R}^{r \times r}$ such that $V_0 = U_0 T_0$. Then, there exists a one-parameter family of linear transformations $T(t)$ such that $V(t) = U(t)T(t)$ for all t and $T(t)$ satisfies the differential equation*

$$\dot{T} = L_r(t)T, \quad T(t_0) = T_0, \qquad (17)$$

where $L_r = U^T L U$ is the orthogonal projection of L to the linear subspace spanned by the columns of U.

Proof. See Theorem 2.4 in Ref. 2. $\square$

Based on this equivalence, it is sufficient to study the behavior of the subspace evolved under the time-dependent linearized dynamics (3) in order to understand the properties of the OTD subspace. Specifically, we use this lemma to prove the main result of this paper, namely, that under a spectral gap condition, the OTD subspace will converge to the dominant directions of the Cauchy–Green strain tensor.

Theorem 3.2. *Let $\xi_i(t, t_0, z_0)$ and $\eta_i(t, t_0, z_0)$ $(i = 1, 2, ..., n)$ denote the eigenvectors of the right and left Cauchy–Green strain tensors, respectively, with corresponding eigenvalues $\lambda_1(t, t_0, z_0) \geq \lambda_2(t, t_0, z_0) \geq \cdots \geq \lambda_n(t, t_0, z_0)$. Moreover, let $U(t) \in \mathbb{R}^{n \times r}$ solve the OTD equation (15) along the trajectory $z(t; t_0, z_0)$. Assume that*

(i) *The subspace U(t) is initiated so that each basis element $u_i(t_0)$ has a non-zero projection on at least one of the eigenvectors $\xi_i(t, t_0, z_0)$, $i = 1, 2, ..., r$ for all times t.*

(ii) *The trajectory $z(t; t_0, z_0)$ is hyperbolic in the sense that there exist constants $a_1 \geq a_2 \geq \cdots \geq a_r > a_{r+1} \geq \cdots \geq a_n$ and $K_i > 0$ ($i = 1, 2, \ldots, n$) such that*

$$\lim_{t \to \infty} \frac{\lambda_i(t, t_0, z_0)}{e^{a_i t}} = K_i. \tag{18}$$

Then, the subspace $U(t)$ aligns with the r most dominant left Cauchy–Green strain eigenvectors, $\eta_i(t, t_0, z_0)$, $i = 1, \ldots, r$, exponentially fast as $t \to \infty$.

Proof. Based on Lemma 3.1 the subspace spanned by the OTD modes is equivalent to the subspace that we obtain if we evolve $V(t) = [v_1(t), \ldots, v_r(t)]$ using the linearized dynamics (3). We represent the initial condition for the equation of variations (3) as $V(t_0) = [v_{0_1}, v_{0_2}, \ldots, v_{0_r}]$ and express it in terms of the orthonormal strain eigenbasis $\{\xi_i(t, t_0, z_0)\}$ as

$$v_{0_j} = \sum_{i=1}^{n} \langle v_{0_j}, \xi_i(t, t_0, z_0) \rangle \xi_i(t, t_0, z_0), \quad j = 1, \ldots, r. \tag{19}$$

Then, Eqs. (4) and (7) applied to Eq. (19) yield

$$v_j(t, t_0, z_0) = \sum_{i=1}^{n} \sqrt{\lambda_i(t, t_0, z_0)} \langle v_{0_j}, \xi_i(t, t_0, z_0) \rangle \eta_i(t, t_0, z_0),$$

$$j = 1, \ldots, r.$$

We then have

$$\langle v_j(t, t_0, z_0), \eta_k(t, t_0, z_0) \rangle = \sqrt{\lambda_k(t, t_0, z_0)} \langle v_{0_j}, \xi_k(t, t_0, z_0) \rangle,$$

$$j, k = 1, \ldots, r.$$

Moreover,

$$\langle v_j(t, t_0, z_0), v_j(t, t_0, z_0) \rangle = \sum_{i=1}^{n} \lambda_i(t, t_0, z_0) \langle v_{0_j}, \xi_i(t, t_0, z_0) \rangle^2,$$

$$j = 1, \ldots, r.$$

Based on this last equation, the cosine of the angle $\alpha_{v_j, \eta_k}(t, t_0, z_0)$ between the vector $v_j(t, t_0, z_0)$ and the eigenvector $\eta_k(t, t_0, z_0)$ of $B_{t_0}^t(z_0)$ obeys the relation

$$\cos \alpha_{v_j, \eta_k} = \frac{\langle v_j(t, t_0, z_0), \eta_k \rangle}{||v_j(t, t_0, z_0)||} = \frac{\sqrt{\lambda_k} \langle v_{0_j}, \xi_k \rangle}{\sqrt{\sum_{i=1}^{n} \lambda_i \langle v_{0_j}, \xi_i \rangle^2}}. \tag{20}$$

Identity (20) and Definition 3.2 yield

$$r\gamma_{v, \eta}^2 = \sum_{j=1}^{r} \sum_{k=1}^{r} \frac{\lambda_k \langle v_{0_j}, \xi_k \rangle^2}{\sum_{i=1}^{n} \lambda_i \langle v_{0_j}, \xi_i \rangle^2}$$

$$= \sum_{j=1}^{r} \frac{\sum_{k=1}^{r} \lambda_k \langle v_{0_j}, \xi_k \rangle^2}{\sum_{i=1}^{n} \lambda_i \langle v_{0_j}, \xi_i \rangle^2} = \sum_{j=1}^{r} \frac{A_j}{A_j + B_j}. \tag{21}$$

where

$$A_j = \sum_{k=1}^{r} \lambda_k \langle v_{0_j}, \xi_k \rangle^2 \geq \lambda_r \sum_{k=1}^{r} \langle v_{0_j}, \xi_k \rangle^2,$$

$$B_j = \sum_{i=1}^{n} \lambda_i \langle v_{0_j}, \xi_i \rangle^2 \leq \lambda_{r+1} \sum_{i=r+1}^{n} \cdot \langle v_{0_j}, \xi_i \rangle^2$$

Note that assumption (ii) implies $\lim_{t \to \infty} \lambda_{r+1}(t, t_0, z_0) / \lambda_r(t, t_0, z_0) = 0$ since

$$\lim_{t \to \infty} \frac{\lambda_{r+1}(t, t_0, z_0)}{\lambda_r(t, t_0, z_0)} = \lim_{t \to \infty} \frac{\lambda_{r+1}(t, t_0, z_0)/e^{a_{r+1} t}}{\lambda_r(t, t_0, z_0)/e^{a_r t}} \frac{e^{a_{r+1} t}}{e^{a_r t}},$$

$$= \frac{K_{r+1}}{K_r} \lim_{t \to \infty} e^{(a_{r+1} - a_r) t}.$$

As a consequence, we obtain from the last inequality the estimate

$$r\gamma_{v, \eta}^2 = \sum_{j=1}^{r} \frac{1}{1 + \frac{B_j}{A_j}} \geq \sum_{j=1}^{r} \frac{1}{1 + \frac{\lambda_{r+1}}{\lambda_r} \frac{\sum_{i=r+1}^{n} \langle v_{0_j}, \xi_i \rangle^2}{\sum_{k=1}^{r} \langle v_{0_j}, \xi_k \rangle^2}}. \tag{22}$$

Note that $\sum_{k=1}^{r} \langle v_{0_j}, \xi_k \rangle^2 > 0$ from the first assumption. Therefore, for the asymptotic limit $t \to \infty$, we have

$$\gamma_{v, \eta}^2 \geq 1.$$

Note that the way we have defined the two subspaces satisfy the assumptions of Lemma 3.1 and therefore we always have $\gamma_{v, \eta}^2 \leq 1$. Thus, $\gamma_{v, \eta}^2 = 1$ which implies that the two subspaces are equivalent. This completes the proof. $\square$

We point out that the hyperbolicity condition (ii) is not a necessary condition. In fact, as long as the ratio $\lambda_{r+1}(t)/\lambda_r(t)$ tends to zero asymptotically, the alignment takes place [see Eq. (22)].

IV. EIGENVALUE CROSSING AND OTD MODES

Theorem 3.2 shows that under appropriate assumptions the OTD modes converge exponentially fast to the eigenspace of the left Cauchy–Green strain tensor associated with the largest Lyapunov exponents. This convergence, however, may be interrupted when the smallest eigenvalue spanned by the OTD modes crosses with the one that is not being spanned. In Fig. 2, we illustrate this situation. The green
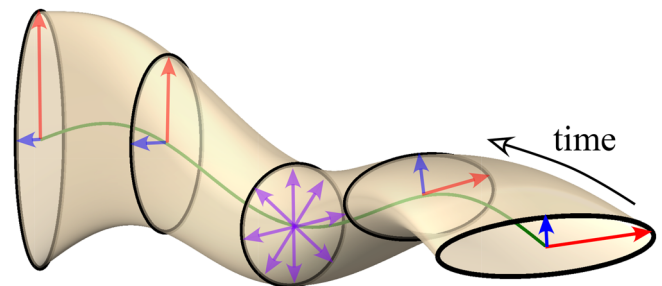

FIG. 2. Evolution of the principal eigenvector (red) of the Cauchy–Green tensor. When eigenvalue crossing occurs, the rate of change becomes unbounded. This is not a problem if all the eigenvectors are resolved but it is important if we evolve only the dominant eigenvectors as it is the case with the OTD modes.

curve denotes the trajectory of the system, while the ellipses indicate the principal axes of the left Cauchy–Green strain tensor. Assuming that we are resolving only one OTD mode, after sufficient time this must have converged to the correct principal direction (red arrow). As we go through the critical time instant where the ellipsoid becomes a circle due to eigenvalue crossing, we observe that the principal directions are instantaneously ill-defined. Immediately after the eigenvalue crossing, the dominant principal direction undergoes a 90° internal rotation compared to immediately before the crossing. The OTD modes, on the other hand, evolve smoothly and are unable to capture such an instantaneously unbounded transition.

In the context of shearless Lagrangian coherent structures, these eigenvalue crossings are referred to as the *Cauchy–Green singularities* and play an important role in the computation of jet cores.[12] In a more general application, the significance of these singularities is pointed out by Lancaster[24] who derived the rate of change of the eigenvectors of an arbitrary symmetric matrix, proving the following theorem.

Theorem 4.1. *The rate of change of the eigenvectors $R(t) \in \mathbb{R}^{n \times n}$ of a symmetric matrix $G(t) \in \mathbb{R}^{n \times n}$ is described by the equation*

$$\dot{R} = RK, \tag{23}$$

where $K(t)$ is a skew-symmetric matrix given by

$$K_{ij}(t) = \frac{\tilde{G}_{ij}(t) + \tilde{G}_{ji}(t)}{2\left(\lambda_j(t) - \lambda_i(t)\right)}, \quad i \neq j, \tag{24}$$

with $\tilde{G}(t) = R^T(t)\dot{G}(t)R(t)$ and $\lambda_i(t)$ being the eigenvalues of $G(t)$.

Proof. See Appendix B or Ref. 24.   □

As λ_i approaches λ_j, $\dot{R}$ becomes unbounded, consistent with the description given previously. This is not a problem if both of the eigen-directions corresponding to λ_i and λ_j are being spanned/resolved by the OTD subspace. The corresponding OTD subspace will just evolve smoothly in this case. However, if the eigenvalue crossing occurs with an eigendirection that is not being spanned by the OTD subspace, the latter will have to instantaneously "steer" towards a direction that is orthogonal to it.

This is not possible since it can be easily seen that for any bounded operator L, the rate of change of the corresponding OTD modes [Eq. (15)] is bounded

$$\dot{U} = QLU \Rightarrow ||\dot{U}|| \leq ||Q|| ||L|| ||U||, \tag{25}$$

where Q is a projection operator (therefore bounded) and U is also bounded. If the eigenvalue crossing involves modes already contained in U, then there is no need for infinite rate of change of the modes, because existing modes just exchange roles (or indices). However, if there is eigenvalue crossing between a mode that is contained in U and another one that is not contained then it is essential to have an evolution of the direction in U to an orthogonal direction (not contained in U) and this cannot happen instantaneously.

Therefore, we see that the convergence of the OTD modes can be compromised if there are eigenvalue crossings involving directions not spanned by the OTD subspace. Note however that such eigenvalue crossing will only influence the convergence of the OTD mode associated with the smallest finite-time Lyapunov exponent captured within the OTD subspace. Assuming that there are no subsequent eigenvalue crossings with larger eigenvalues, one expects that the OTD subspace would capture the directions associated with the largest finite-time Lyapunov exponents.

Thus, for systems where eigenvalue crossing occurs, a larger OTD subspace guarantees that the principal axes associated with the largest Lyapunov exponents will be captured by the OTD modes. Note that *such convergence does not depend on the physical dimension of the system or the intrinsic dimension of the underlying attractor*. This is particularly important for systems where instabilities of isolated modes emerge out of high-dimensional stochastic backgrounds, such as intermittent dissipation bursts in turbulence[16] or focusing instabilities occurring in nonlinear stochastic waves.[9,16]

As we demonstrate in Sec. V, for many practical applications, it is sufficient to compute just the maximum finite-time Lyapunov exponent so what is important is to have the most dominant direction of the Cauchy–Green tensor captured by the OTD subspace. The benefits of such an approach are, of course, more important as the dimension n of the underlying system increases.

V. REDUCED-ORDER COMPUTATION OF FINITE-TIME LYAPUNOV EXPONENTS

As shown above, the OTD modes are able to extract the transiently most unstable directions independent of the attractor dimensionality. Consequently, the OTD basis can be used for the reduced-order approximation of finite-time Lyapunov exponents (FLTE) of high-dimensional systems. Specifically, we apply the following approximation scheme for the computation of a FTLE corresponding to a time interval of length T.

(1) Advect the trajectory $z(t; t_0, z_0)$ for an interval $t \in [t_0, t_0 + T]$ where z_0 is the initial point.
(2) Compute the r-dimensional OTD subspace $U(t)$ corresponding to this trajectory.
(3) Compute the low-dimensional fundamental solution matrix $\Phi_{t_0}^t \in \mathbb{R}^{r \times r}$ using the reduced linear dynamical system

$$\frac{d}{dt}\Phi_{t_0}^t = L_r(z(t), t)\Phi_{t_0}^t, \quad \Phi_{t_0}^{t_0} = I, \quad t \in [t_0, t_0 + T], \tag{26}$$

where $L_r(z, t) = U(t)^T L(z, t) U(t)$ is the projection of the full linearized operator L onto the OTD subspace, which contains the most unstable directions.
(4) Compute the reduced-order finite-time Lyapunov exponents,

$$\Gamma_i(t_0 + T, t_0, z_0) = \frac{1}{T} \log \sqrt{\gamma_i(t_0 + T, t_0, z_0)}, \quad i = 1, \ldots, r, \tag{27}$$

where γ_i denotes the eigenvalues of the reduced-order right Cauchy–Green strain tensor $(\Phi_{t_0}^{t_0+T})^{\top}\Phi_{t_0}^{t_0+T}$, ordered such that $\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_n$.

The described algorithm allows for the computation of FTLE by solving for each initial point, z_0, one n-dimensional equation, as well as additional r equations which are n-dimensional for the OTD subspace. This cost has to be summed together with the solution of the reduced-order linear system. To this end, we have the total cost for the reduced-order and full algorithms:

- Full-order calculation cost: $(n + 1) \times n$ consisting of $n + 1$ equations of dimension n as follows:

$$\dot{z} = f(z, t) \quad (n \text{ equations}), \tag{28a}$$

$$\frac{d}{dt}\nabla F_{t_0}^t(z_0) = L(z(t), t)\nabla F_{t_0}^t(z_0) \quad (n \times n \text{ equations}). \tag{28b}$$

- Reduced-order calculation cost: $n \times (r + 1) + r^2$ consisting of $r + 1$ equations of dimension n and r equations of dimension r as follows:

$$\dot{z} = f(z, t) \quad (n \text{ equations}), \tag{29a}$$

$$\dot{U} = L(z(t), t)U - U(U^T L(z(t), t)U) \quad (n \times r \text{ equations}), \tag{29b}$$

$$\frac{d}{dt}\Phi_{t_0}^t = L_r(z(t), t)\Phi_{t_0}^t \quad (r \times r \text{ equations}). \tag{29c}$$

We observe that the cost of computing FTLE using the full system is $\mathcal{O}(n^2)$. The computational cost of the reduced-order FTLE, however, is $\mathcal{O}(n)$ for fixed r. In principle, the dimension r of the reduced system can be as low as 1. But, as we will see in what follows, a one-dimensional OTD subspace may lead to *false troughs* in the FTLE field. A higher dimensional OTD subspace returns more accurate estimation of the FTLE field and often circumvents the false troughs.

Figure 3 shows the computational cost of the full FTLE calculation versus the reduced-order FTLE calculation. Note that for low-dimensional systems (small n) the reduction in the computational cost is not significant. However, as the system dimension increases ($n \gg r$), the computational cost of the reduced-order FTLE can be orders of magnitude lower than the full FTLE computations.

A. The ABC flow

As the first example, we consider the ABC (Arnold–Beltrami–Childress) flow with three dimensional velocity field

$$\begin{aligned}\dot{z}_1 &= A\sin(z_3) + C\cos(z_2),\\ \dot{z}_2 &= B\sin(z_1) + A\cos(z_3),\\ \dot{z}_3 &= C\sin(z_2) + B\cos(z_1),\end{aligned} \tag{30}$$

which is an exact solution of the Euler equation for the ideal incompressible fluids.[1] The ABC flow has served as a prototype example for testing numerical methods for computing Lagrangian coherent structures.[14,15,19] Here, we set $A = \sqrt{3}$,
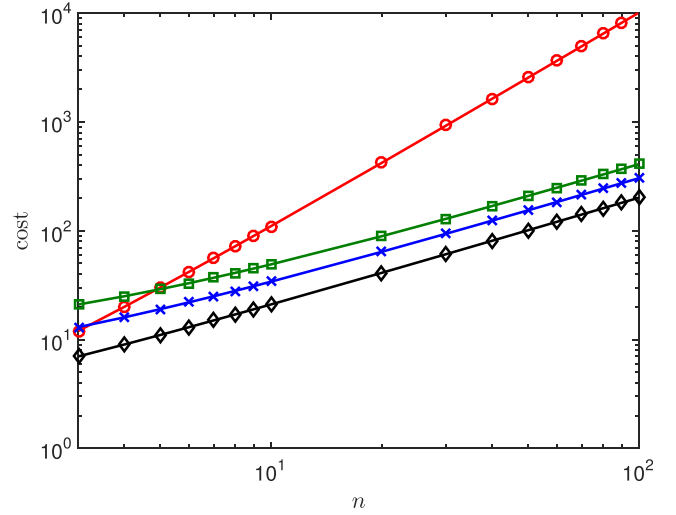

FIG. 3. The computational cost of the FTLE using the full system (28) (red circles), and the reduced system (29) with $r = 1$ (black diamonds), $r = 2$ (blue crosses), and $r = 3$ (green squares). The integer n is the system dimension. The computational cost is measured as the number of scalar ODEs that need to be solved for each case.

$B = \sqrt{2}$, and $C = 1$ which allows for the existence of chaotic Lagrangian trajectories.[19]

We compute the FTLE on a 251×251 grid of initial conditions on each face of the cube $[0, 2\pi]^3$ with an integration time of length $T = 8$. We use finite differences to approximate the deformation gradient $\nabla F_{t_0}^{t_0+T}$. To increase the finite difference accuracy, we use six auxiliary grids ($z_{0,1} \pm h, z_{0,2} \pm h, z_{0,3} \pm h$) around each grid point $z_0 = (z_{0,1}, z_{0,2}, z_{0,3})$. The parameter h controls the accuracy of the finite differences and is set to $h = 10^{-8}$ in the following. We refer to Refs. 13 and 25 for further details on the approximation of the deformation gradient. Once the deformation gradient is approximated, computing the Cauchy–Green strain tensor and consequently the FTLE is straightforward. Note that the deformation gradient can also be obtained by numerically integrating the equation of variations (3).

To compute the reduced-order FTLE, we numerically integrate system (29) from the initial grid points z_0 with an integration time of length T. The initial condition $U(0)$ for the OTD equation (29b) is the r most dominant eigenvectors of the symmetric linear operator at $t = 0$, i.e., $L^s = (L(z_0, 0) + L(z_0, 0)^T)/2$. This choice of the OTD initial condition is made since these eigenvectors are the instantaneously most unstable directions.[21]

Figure 4 shows the reduced-order FTLE fields using OTD reduction of sizes $r = 1$ (left column) and $r = 2$ (middle column). The true FTLE field (right column) is also shown for comparison. Each row of the figure shows a different cross section of the field. For $r = 2$, the reduced-order FTLE is close to the true FTLE field. This is also true qualitatively for $r = 1$. However, for $r = 1$, the reduced FTLE exhibits additional troughs that do not exist in the true FTLE field.

We demonstrate in Fig. 5 that these false troughs occur when there is a crossing (or near crossing) between the first and second most dominant eigenvalues of the right Cauchy–Green tensor, i.e., λ_1 and λ_2. Recall that this eigenvalue crossing violates condition (ii) of Theorem 3.2 and hence, the discrepancy
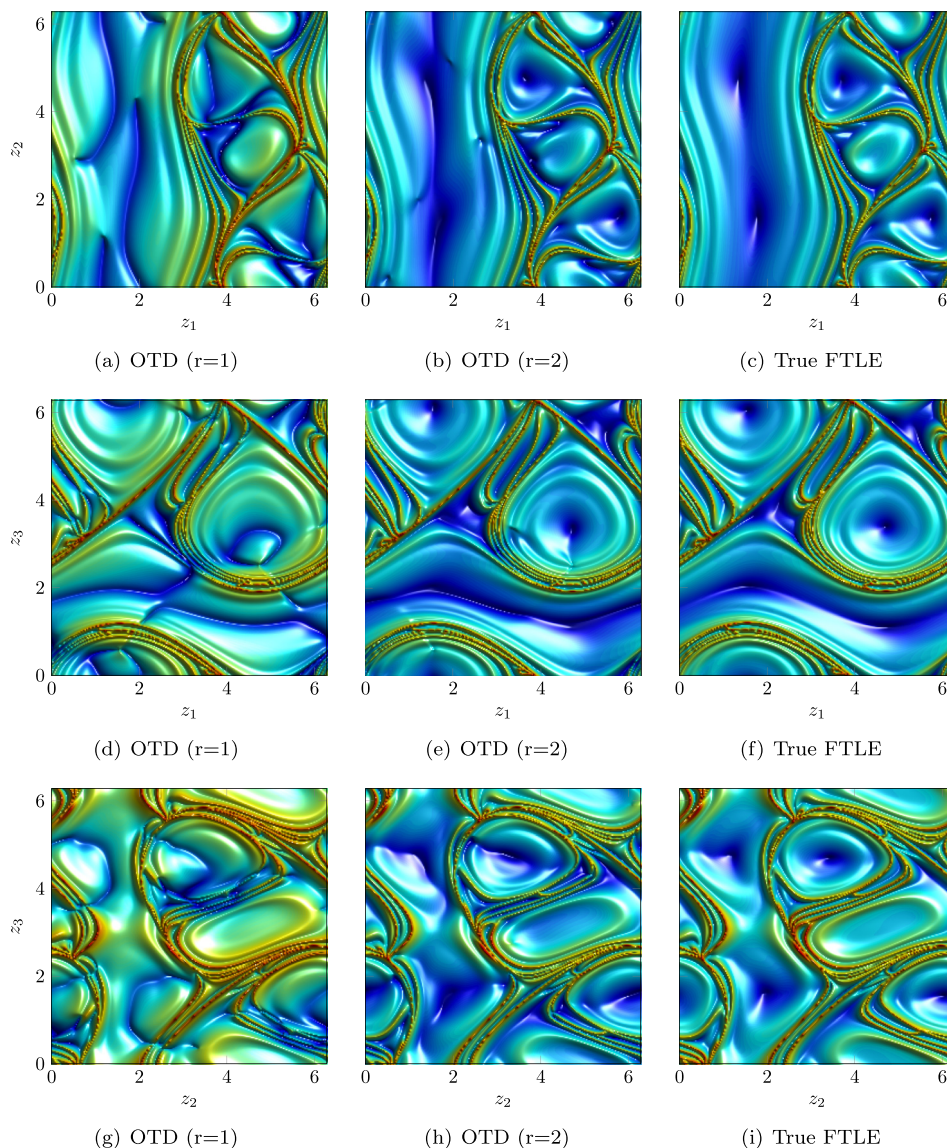
FIG. 4. ABC flow: comparison of the true FTLE (right column) with OTD reduction of size $r = 1$ (left column) and $r = 2$ (middle column).

between the reduced-order FTLE and the true FTLE is expected. The false troughs are eliminated as we increase the dimension of the OTD reduction from $r = 1$ to $r = 2$. This is due to the fact that there is no eigenvalue crossing between λ_2 and λ_3.

The above observation is demonstrated in Fig. 5 where the reduced-order FTLEs at two select points are compared against the true FTLE as the integration time T varies. In Fig. 5(a), the initial point $z_0 = (2.0, 0.6, 0.0)$ is chosen to be at one of the false troughs. An eigenvalue crossing occurs at approximately $T = 0.2$. For advection times before $T = 0.2$, both one- and two-dimensional reductions follow the largest FTLE very closely. However, near the eigenvalue crossing the most dominant eigenvector of the Cauchy–Green strain tensor undergoes rotation with an unbounded rate. As demonstrated in inequality 25, the one-dimensional OTD reduction cannot follow such rapid rotation. This is confirmed in Fig. 5(a) where it can be seen that the FTLE obtained from the one-dimensional OTD reduction, after $T = 0.2$, follows the second most dominant FTLE instead of undergoing a rapid change that is required to follow the most dominant FTLE. Near $T = 1$, the FTLE of the one-dimensional reduction

departs from non-dominant FTLEs and starts increasing. This behavior is to be expected since in the remaining advection time no eignevalue crossing occurs and as a result the one-dimensional reduction continues to recover to converge to the most dominant eigenvector—however slowly.

The two-dimensional OTD reduction, however, recovers much faster compared to the one-dimensional OTD reduction after $T = 0.2$. We observe that the two eigenvalues of OTD ($r = 2$) approach each other nearly at $T = 0.5$ causing in a relatively rapid repulsion of these two eigen-directions. This repulsion results in a rapid rotation of the dominant OTD vector towards the dominant eigen-direction of the Cauchy–Green tensor. This mechanism of rapid correction of the OTD vectors is entirely absent in the one-dimensional OTD reduction.

Figure 5(b) shows the evolution of the same quantities along a different ABC trajectory. Here, eigenvalue crossings are absent. As a result, both one-dimensional and two-dimensional OTD reductions closely follow the most dominant FTLE and in the case of $r = 2$ both most and second most dominant FTLEs. This demonstrates that, in the absence of eigenvalue crossing, one-dimensional OTD reduction accurately approximates the most dominant FTLE (Fig. 6).

(a) $z_0 = (2.0, 0.6, 0.0)$
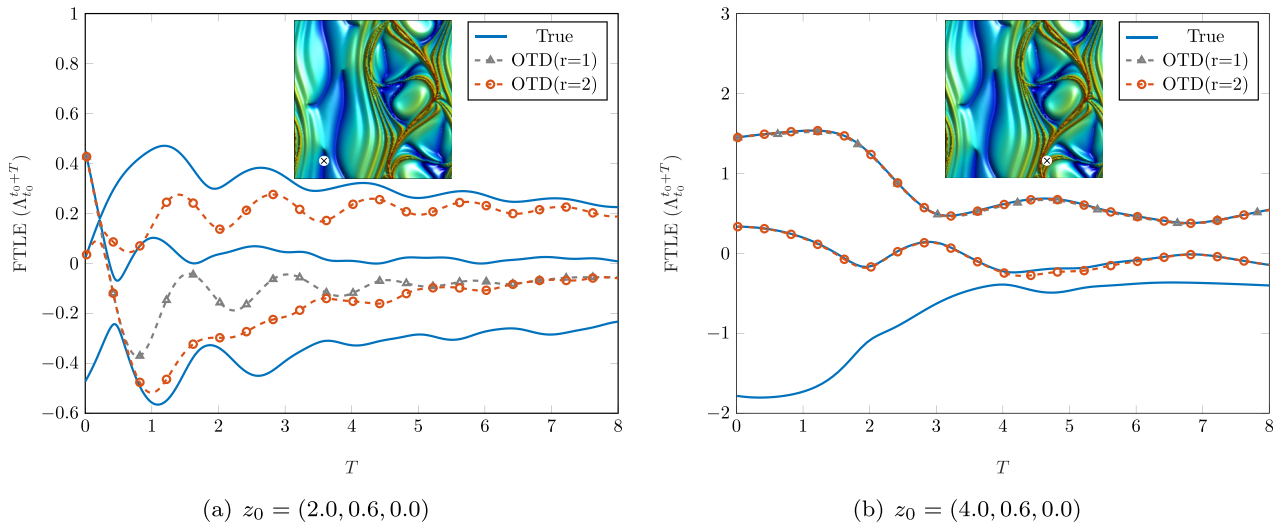

(b) $z_0 = (4.0, 0.6, 0.0)$

FIG. 5. ABC flow: FTLE versus advection time calculated with full dynamics (solid blue), one-dimensional OTD reduction (dashed gray line with triangle symbols) and two-dimensional OTD reduction (red line with circle symbols) with initial points of: (a) $z_0 = (2.0, 0.6, 0.0)$, and (b) $z_0 = (4.0, 0.6, 0.0)$. The initial points are shown with a cross symbol in the $z_1 - z_2$ plane.
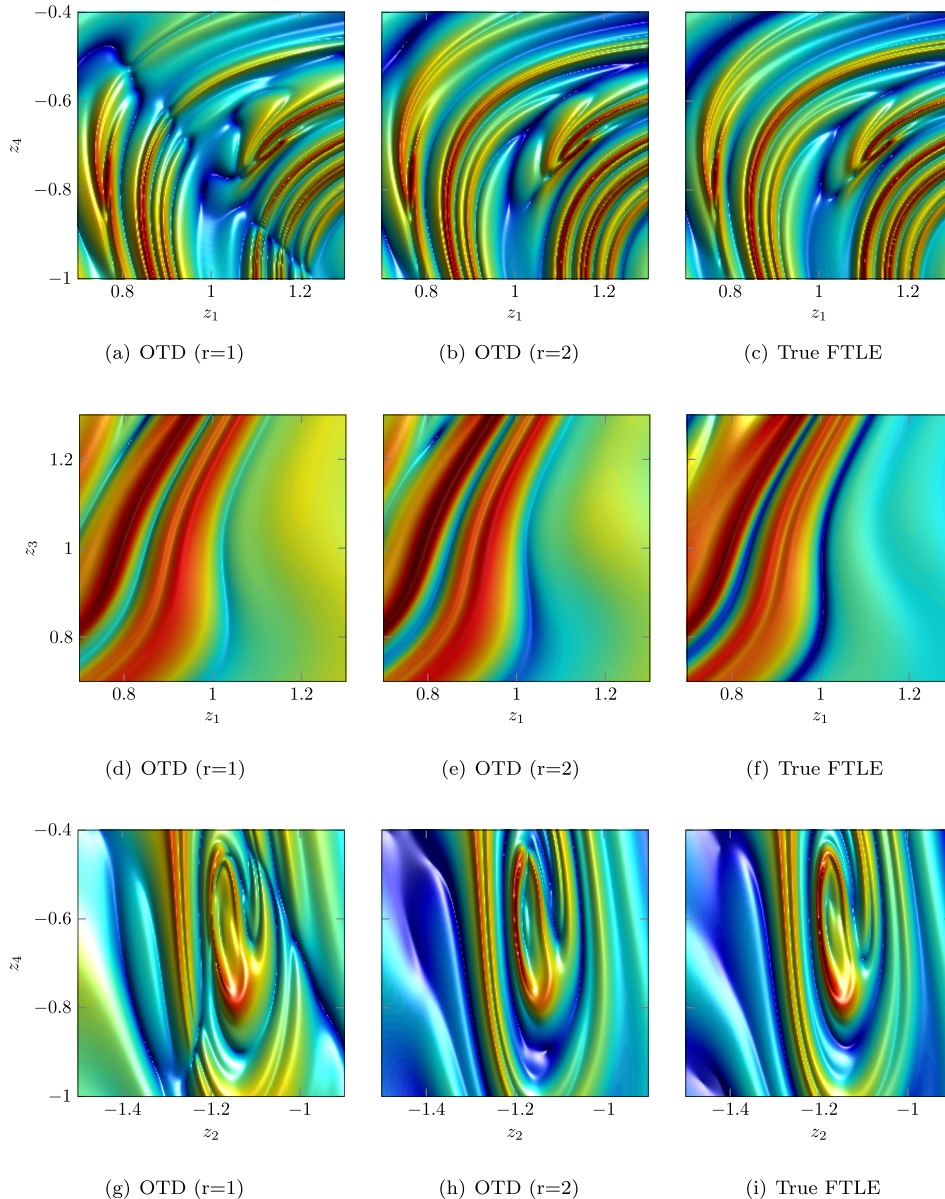

(a) OTD (r=1)

(b) OTD (r=2)

(c) True FTLE

(d) OTD (r=1)

(e) OTD (r=2)

(f) True FTLE

FIG. 6. Six-dimensional Charney–DeVore model: comparison of the true FTLE (rightmost column) with OTD reduction of size $r = 1$ (leftmost column) and $r = 2$ middle column. Each row shows results of a section of the phase space.

(g) OTD (r=1)

(h) OTD (r=2)

(i) True FTLE

B. Charney–DeVore model with regime transitions

In this section, we apply the OTD reduction to the six-dimensional truncation of the equations for barotropic flow in a plane channel with orography that are known as the Charney–DeVore (CDV) model. The truncated system is given by

$$
\begin{aligned}
\dot{z}_1 &= \gamma_1^* z_3 - C(z_1 - z_1^*),\\
\dot{z}_2 &= -(\alpha_1 z_1 - \beta_1)z_3 - C z_2 - \delta_1 z_4 z_6,\\
\dot{z}_3 &= (\alpha_1 z_1 - \beta_1)z_2 - \gamma_1 z_1 - C z_3 + \delta_1 z_4 z_5,\\
\dot{z}_4 &= \gamma_2^* z_6 - C(c_4 - z_4^*) + \epsilon(z_2 z_6 - z_3 z_5),\\
\dot{z}_5 &= -(\alpha_2 z_1 - \beta_2)z_6 - C z_5 - \delta_2 z_4 z_3,\\
\dot{z}_6 &= (\alpha_2 z_1 - \beta_2)z_2 - \gamma_2 z_4 - C z_6 + \delta_2 z_4 z_2.
\end{aligned}
\tag{31}
$$

The model coefficients are given by

$$
\begin{aligned}
\alpha_m &= \frac{8\sqrt{2}}{\pi}\frac{m^2}{4m^2-1}\frac{b^2+m^2-1}{b^2+m^2}, \quad \beta_m = \frac{\beta b^2}{b^2+m^2},\\
\delta_m &= \frac{64\sqrt{2}}{15\pi}\frac{b^2-m^2+1}{b^2+m^2}, \quad \gamma_m^* = \gamma\frac{4m}{4m^2-1}\frac{\sqrt{2}b}{\pi},\\
\epsilon &= \frac{16\sqrt{2}}{5\pi}, \quad \gamma_m = \gamma\frac{4m^3}{4m^2-1}\frac{\sqrt{2}b}{\pi(b^2+m^2)}.
\end{aligned}
\tag{32}
$$

Following Ref. 11, we set $(z_1^*, z_4^*, C, \beta, \gamma, b) = (0.95, -0.76095, 0.1, 1.25, 0.2, 0.5)$. For these parameters, the CDV model generates regime transitions due to the interaction of barotropic and topographic instabilities. It was shown in Ref. 10 that a Proper Orthogonal Decomposition (POD) reduction of this model to three leading POD modes resolves 97% of cumulative variance but cannot capture the chaotic regime transitions present in the six-dimensional model. These highly transient instabilities render this model an appropriate test case for evaluating the performance of the OTD reduction in computing the FTLE. Moreover, using this model we assess the performance of the OTD reduction for a higher dimensional problem.

We compute the true FTLE using the finite difference method analogous to the ABC flow problem. Similarly, the initial condition $U(0)$ for the OTD equation (29b) is the r most dominant eigenvectors of the symmetric linearized operator. The fourth-order Runge-Kutta scheme with time step size $\Delta t = 0.4$ (days) is used for the numerical integration of Eqs. (28) and (29). Using the smaller time step $\Delta t = 0.2$ (days) did not change the numerical results.

Figure 6 shows the reduced FTLE fields obtained from the one-dimensional reduction (left column) and two-dimensional reduction (middle column) as well as the full FTLE (right column). Each row shows a two-dimensional cross section of the phase space. As in the ABC flow, the reduced FTLE fields agree qualitatively with the full FTLE field. However, for $r=1$ reduction, some false troughs are observed. Away from the false troughs, the one-dimensional reduction estimates the FTLE accurately. For $r=2$ reduction, there are no false troughs.

In Fig. 7, the FTLE values are plotted as a function of the integration time. These correspond to the point $z_0 = (1.14, 0, 0, -0.91, 0, 0)$, marked by a cross symbol in
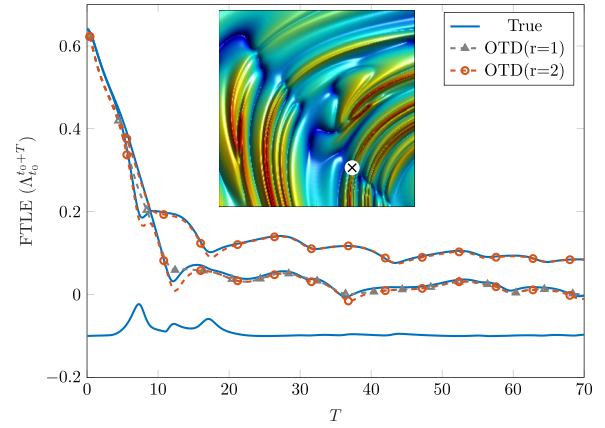

FIG. 7. Six-dimensional Charney–DeVore model: FTLE versus advection time calculated with full dynamics (solid blue), one-dimensional OTD reduction (dashed gray line with triangle symbols), and two-dimensional OTD reduction (red line with circle symbols) with initial points of $z_{0,1} = 1.14$, $z_{0,4} = -0.91$ and other coordinates being zero. The initial point is shown with a cross symbol in the $z_1 - z_4$ plane.

the inset, which lies on a false trough in the $z_1 - z_4$ section. An eigenvalue crossing occurs at approximately $T = 9$. Analogous to the ABC flow, we observe that before $T = 9$ both one- and two-dimensional OTD reductions capture the most dominant eigenvalues of the Cauchy–Green strain tensor. After $T = 9$, the one-dimensional OTD reduction continuously follows the second most dominant eigen-direction of the Cauchy–Green strain tensor. The inability of a single OTD vector to undergo a dramatic rotation to follow the most dominant eigen-direction of the Cauchy–Green tensor leads to the false trough that is observed in Fig. 6(a). The two-dimensional OTD subspace, however, converges to the subspace spanned by the two most dominant eigenvectors of the Cauchy–Green strain tensor. This convergence is guaranteed by Theorem 3.2 since the second and the third Cauchy–Green eigenvalues (λ_2, λ_3) do not cross (see Fig. 7). As a result, the two-dimensional $(r = 2)$ OTD reduction closely approximates the two large FTLEs for all times.

The above results demonstrate that, as long as λ_2 and λ_3 do not coincide, a reduced FTLE with two OTD modes reliably approximates the dominant FTLEs of the full system regardless of the system's dimension n. This amounts to a significant reduction in the computational cost of the FTLE evaluations, and the gain in the computational speed is larger as the dimension of the dynamical system increases. Investigating the numerical performance of the OTD reduction for infinite-dimensional systems is the subject of future study.

VI. CONCLUSIONS

We have examined the properties of the optimally time-dependent (OTD) modes for general time-dependent dynamical systems. Specifically, we have shown that under mild conditions, related to the spectrum of the Cauchy–Green tensor, the *OTD modes converge exponentially fast to the eigendirections of the Cauchy–Green tensor associated with the dominant eigenvalues (largest finite-time Lyapunov exponents). Therefore, the OTD modes can be employed to*

extract time-dependent subspaces that encode information associated with the transient dynamics and accurately characterize the associated finite-time instabilities. This is an important development relative to known shortcomings of proper orthogonal decomposition models, which may change or even invert stability characteristics.[3,29]

We have applied the derived result on the formulation of a reduced order algorithm for the computation of the maximum finite-time Lyapunov exponent, a measure that has been used for the quantification of Lagrangian Coherent Structures. We demonstrated the derived results through two specific examples, the three-dimensional ABC flow and the six-dimensional Charney-DeVore model. In both cases, we thoroughly analyzed the limitations due to dimensionality reduction in combination with eigenvalue crossing. Apart from its value as a computational method for finite-time Lyapunov exponents, the presented result paves the way for the development of efficient control and prediction strategies for chaotic dynamical systems exhibiting transient features, such as extreme events and off-equilibrium dynamics.

ACKNOWLEDGMENTS

APPENDIX A: THE OTD MODES AND THE DYNAMICALLY ORTHOGONAL MODES

The set of OTD equations have the same form as the Dynamically Orthogonal (DO) field equations.[31,33] In fact, because the minimization of the function $\mathcal{F}$ is performed only over the rate of change of the basis elements, $\dot{u}_i(t)$, [and not the basis elements $u_i(t)$] one can follow exactly the same steps as in Theorem 2.1 in Ref. 2 to show that the evolution equations (15) can be derived for the general case of a nonlinear operator $\mathcal{N}(U)$. More specifically, by minimization of the functional

$$\mathcal{F}(\dot{u}_1, \dot{u}_2, ..., \dot{u}_r) = \sum_{i=1}^{r} \left\| \frac{\partial u_i(t)}{\partial t} - \mathcal{N}(z(t), U(t), t) \right\|^2, \quad \text{(A1)}$$

we can obtain the evolution equations.

$$\frac{\partial U}{\partial t} = \mathcal{N}(U) - UU^T \mathcal{N}(U). \quad \text{(A2)}$$

For the case of the DO equations, the nonlinear operator takes the form

$$\mathcal{N}(U) = E^\omega \left[\mathcal{L}(\bar{u} + \sum_{i=1}^{r} u_i Y_i) Y_j \right] C_{Y_i Y_j}^{-1}, \quad \text{(A3)}$$

where we have used the notation in Ref. 31, i.e., $\mathcal{L}$ is the right hand side of the stochastic partial differential equation (PDE), $\bar{u}$ is the trajectory of the mean, $Y_i(t)$ are the stochastic coefficients and $C_{Y_i Y_j}(t)$ is their covariance matrix. The

operation E^ω denotes the average with respect to an appropriate probability measure. Conversely, one can obtain the OTD mode equations from the DO equations simply by considering their deterministic limit, i.e., by taking the limit $C_{Y_i Y_j}(t) \to 0$.

APPENDIX B: PROOF OF THEOREM 4.1

We consider the rate of change of the eigenvectors of a general one-parameter family of symmetric matrices, $G(t) \in \mathbb{R}^{n \times n}$. For such a matrix, we have the eigenvalue problem

$$G(t)R(t) = R(t)\Lambda(t), \quad \text{(B1)}$$

where the columns of $R(t) \in \mathbb{R}^{n \times n}$ are the eigenvectors of $G(t)$ and $\Lambda(t) \in \mathbb{R}^{n \times n}$ is the diagonal matrix of corresponding eigenvalues. Note that, since G is symmetric, its eigenvectors are orthogonal and therefore, $RR^T = I$. Differentiating with respect to time, we obtain

$$\dot{G}R + G\dot{R} = \dot{R}\Lambda + R\dot{\Lambda}.$$

Multiplying the above equation by R^T from the left yields

$$\dot{\Lambda} = R^T \dot{G} R + \Lambda R^T \dot{R} - R^T \dot{R} \Lambda.$$

Since $R^T R = I$, we have:

$$\dot{R}^T R + R^T \dot{R} = 0.$$

Let $K = R^T \dot{R}$. From the above relation, we observe that K is a skew-symmetric matrix. Denoting $\tilde{G} = R^T \dot{G} R$, we have

$$\dot{\Lambda} = \tilde{G} + \Lambda K - K\Lambda.$$

For off-diagonal terms $i \neq j$, we have $\dot{\Lambda}_{ij} = 0$. Therefore,

$$\tilde{G}_{ij} = \lambda_j K_{ij} - \lambda_i K_{ij}, \quad i \neq j.$$

Transposing the above relation

$$\tilde{G}_{ji} = \lambda_i K_{ji} - \lambda_j K_{ji}, \quad i \neq j.$$

Using the skew-symmetric property of K, i.e., $K_{ij} = -K_{ji}$, and summing up the above two equations yields

$$K_{ij} = \frac{\tilde{G}_{ij} + \tilde{G}_{ji}}{2(\lambda_j - \lambda_i)}, \quad i \neq j. \quad \text{(B2)}$$

This results in the following closed-form evolution equations for Λ and R,

$$\dot{\Lambda} = \text{diag}(\tilde{G}), \quad \text{(B3)}$$

$$\dot{R} = RK. \quad \text{(B4)}$$

[1]V. I. Arnold and B. A. Khesin, *Topological Methods in Hydrodynamics* (Springer, 1998), Vol. 125.

[2]H. Babaee and T. P. Sapsis, "A minimization principle for the description of modes associated with finite-time instabilities," Proc. R. Soc. A **472**, 20150779 (2016).

[3]S. L. Brunton and B. R. Noack, "Closed-loop turbulence control: Progress and challenges," Appl. Mech. Rev. **67**, 050801 (2015).

[4]G. J. Chandler and R. R. Kerswell, "Invariant recurrent solutions embedded in a turbulent two-dimensional Kolmogorov flow," J. Fluid Mech. **722**, 554–595 (2013).

[5]M. Cheng, T. Hou, and Z. Zhang, "A dynamically bi-orthogonal method for time-dependent stochastic PDEs I: Adaptivity and generalizations," J. Comput. Phys. **242**, 753–776 (2013).

[6]M. Cheng, T. Hou, and Z. Zhang, "A dynamically bi-orthogonal method for time-dependent stochastic PDEs I: Derivation and algorithms," J. Comput. Phys. **242**, 843–868 (2013).

[7]S. P. Cornelius, W. L. Kath, and A. E. Motter, "Realistic control of network dynamics," Nat. Commun. **4**, 1942 (2013).

[8]W. Cousins and T. P. Sapsis, "Quantification and prediction of extreme events in a one-dimensional nonlinear dispersive wave model," Physica D **280**, 48–58 (2014).

[9]W. Cousins and T. P. Sapsis, "Reduced order precursors of rare events in unidirectional nonlinear water waves," J. Fluid Mech. **790**, 368–388 (2016).

[10]D. T. Crommelin and A. J. Majda, "Strategies for model reduction: Comparing different optimal bases," J. Atmos. Sci. **61**(17), 2206–2217 (2004).

[11]D. T. Crommelin, J. D. Opsteegh, and F. Verhulst, "A mechanism for atmospheric regime behavior," J. Atmos. Sci. **61**(12), 1406–1419 (2004).

[12]M. Farazmand, D. Blazevski, and G. Haller, "Shearless transport barriers in unsteady two-dimensional flows and maps," Physica D **278–279**, 44–57 (2014).

[13]M. Farazmand and G. Haller, "Computing Lagrangian coherent structures from their variational theory," Chaos **22**, 013128 (2012).

[14]M. Farazmand and G. Haller, "Attracting and repelling Lagrangian coherent structures from a single computation," Chaos **4**, 023101 (2013).

[15]M. Farazmand and G. Haller, "Polar rotation angle identifies elliptic islands in unsteady dynamical systems," Physica D **315**, 1–12 (2016).

[16]M. Farazmand and T. P. Sapsis, "Dynamical indicators for the prediction of bursting phenomena in high-dimensional systems," Phys. Rev. E **94**, 032212 (2016).

[17]M. Farazmand and T. P. Sapsis, "Reduced-order prediction of rogue waves in two-dimensional deep-water waves," J. Comput. Phys. **340**, 418–434 (2017).

[18]F. Ginelli, P. Poggi, A. Turchi, H. Chaté, R. Livi, and A. Politi, "Characterizing dynamics with covariant Lyapunov vectors," Phys. Rev. Lett. **99**(13), 130601 (2007).

[19]G. Haller, "Distinguished material surfaces and coherent structures in 3D fluid flows," Physica D **149**, 248–277 (2001).

[20]G. Haller, "Lagrangian structures and the rate of strain in a partition of two-dimensional turbulence," Phys. Fluids **13**(11), 3365–3385 (2001).

[21]G. Haller and T. Sapsis, "Localized instability and attraction along invariant manifolds," SIAM J. Appl. Dyn. Syst. **9**(2), 611–633 (2010).

[22]G. Haller and G. Yuan, "Lagrangian coherent structures and mixing in two-dimensional turbulence," Physica D **147**, 352–370 (2000).

[23]D. Karrasch, M. Farazmand, and G. Haller, "Attraction-based computation of hyperbolic Lagrangian coherent structures," J. Comput. Dyn. **2**(1), 83–93 (2015).

[24]P. Lancaster, "On eigenvalues of matrices dependent on a parameter," Numer. Math. **6**, 377 (1964).

[25]F. Lekien and S. D. Ross, "The computation of finite-time Lyapunov exponents on unstructured meshes and for non-Euclidean manifolds," Chaos **20**, 017505 (2010).

[26]F. Lekien, S. Shadden, and J. Marsden, "Lagrangian coherent structures in n-dimensional systems," J. Math. Phys. **48**, 065404 (2007).

[27]A. J. Majda, D. Qi, and T. P. Sapsis, "Blended particle filters for large-dimensional chaotic dynamical systems," Proc. Natl. Acad. Sci. **111**(21), 7511–7516 (2014).

[28]E. Musharbash, F. Nobile, and T. Zhou, "Error analysis of the dynamically orthogonal approximation of time dependent random PDEs," SIAM J. Sci. Comput. **37**(2), A776–A810 (2015).

[29]B. R. Noack, K. Afanasiev, M. Morzynski, G. Tadmor, and F. Thiele, "A hierarchy of low-dimensional models for the transient and post-transient cylinder wake," J. Fluid Mech. **497**, 335 (2003).

[30]W. Ott and J. A. Yorke, "When lyapunov exponents fail to exist," Phys. Rev. E **78**(5), 056203 (2008).

[31]T. P. Sapsis, "Attractor local dimensionality, nonlinear energy transfers, and finite-time instabilities in unstable dynamical systems with applications to 2D fluid flows," Proc. R. Soc. A **469**(2153), 20120550 (2013).

[32]T. P. Sapsis and H. A. Dijkstra, "Interaction of noise and nonlinear dynamics in the double-gyre wind-driven ocean circulation," J. Phys. Oceanogr. **43**, 366–381 (2013).

[33]T. P. Sapsis and P. F. J. Lermusiaux, "Dynamically orthogonal field equations for continuous stochastic dynamical systems," Physica D **238**, 2347–2360 (2009).

[34]T. P. Sapsis and P. F. J. Lermusiaux, "Dynamical criteria for the evolution of the stochastic dimensionality in flows with uncertainty," Physica D **241**, 60 (2012).

[35]T. P. Sapsis and A. J. Majda, "A statistically accurate modified quasilinear Gaussian closure for uncertainty quantification in turbulent dynamical systems," Physica D **252**, 34–45 (2013).

[36]T. P. Sapsis and A. J. Majda, "Blending modified Gaussian closure and non-gaussian reduced subspace methods for turbulent dynamical systems," J. Nonlinear Sci. **23**, 1039 (2013).

[37]T. P. Sapsis and A. J. Majda, "Statistically accurate low order models for uncertainty quantification in turbulent," Proc. Natl. Acad. Sci. **110**, 13705–13710 (2013).

[38]Y. Susuki, I. Mezic, and I. Mezić, "Nonlinear Koopman modes and a precursor to power system swing instabilities," IEEE Trans. Power Syst. **27**(3), 1182–1191 (2012).