

## Cross proper orthogonal decomposition

André V. G. Cavalieri \* and André F. C. da Silva †*Divisão de Engenharia Aeroespacial, Instituto Tecnológico de Aeronáutica, São José dos Campos, SP, Brazil*(Received 11 August 2020; accepted 22 December 2020;  
published 11 January 2021)

A method is proposed in order to optimally decompose the trace of cross-covariances of flow fluctuations, such as Reynolds stresses. Such method, referred to as cross proper orthogonal decomposition (CPOD), leads to a basis of modes extracted from a flow database that are optimal in representing an inner product related to the cross-covariance of interest. A sample application is shown for the representation of Reynolds shear stress in a turbulent channel flow with friction Reynolds number equal to 179. Leading modes are shown to comprise streamwise vortices and streaks with phase opposition between streamwise ( $u'$ ) and wall-normal ( $v'$ ) velocities, representing ejections and sweeps, and higher-order modes show similar structures, but with  $u'$  and  $v'$  in phase. A combination of such structures leads to an accurate reconstruction of the Reynolds stress, and consequently of the mean flow, with a reasonable near-wall reconstruction with the leading CPOD mode pair (even and odd modes) for each considered wave number, and a close match of the profiles with the five leading CPOD mode pairs. The present method is thus a valuable modal decomposition technique targeting cross-covariances of flow quantities such as Reynolds stresses.

DOI: [10.1103/PhysRevFluids.6.014602](https://doi.org/10.1103/PhysRevFluids.6.014602)

### I. INTRODUCTION

Coherent structures are an important feature of various turbulent flows, such as free shear layers and jets [1] or wall-bounded turbulence [2]. Despite its complexity, turbulent flow tends to organize into features such as Kelvin-Helmholtz wave packets in free-shear flows, or streamwise vortices and streaks of various length scales in wall turbulence. It is clear that most quantities of practical interest, such as sound radiation, turbulent mixing, wall shear stress, or heat transfer, are a function of the behavior of coherent structures.

Despite their relevance, it is hard to objectively define a coherent structure. The emergence of large databases from experimental and numerical data popularized the application of modal decomposition to study complex turbulent flows [3], and thus obtain coherent structures featuring some relevant property from the fields. A popular method is proper orthogonal decomposition (POD), which decomposes a space-time field into orthogonal spatial functions, POD modes, which optimally represent the turbulent kinetic energy [4]. Other fields often refer to the same technique as Karhunen-Loève decomposition or principal component analysis [5]. POD modes are eigenfunctions of the two-point correlation tensor of the flow, and leading POD modes are spatial structures that optimally represent the fluctuation field. POD has been extensively used to characterize turbulent flows; for wall-bounded turbulence, early efforts can be found in the work by Aubry *et al.* [6] in the extraction and modeling of large-scale structures in a turbulent boundary layer.

\*andre@ita.br

†andref@ita.br

More recent examples involve turbulent pipe flow, with POD of simulation [7] and experimental [8] data showing the organization of coherent structures in streaks of streamwise velocity, sided by streamwise vortices. POD modes form an orthonormal basis, which is useful to obtain reduced-order models using a Galerkin projection of the flow equations onto a basis formed by the leading modes [6,9,10].

More recently, its frequency-domain variant, spectral POD or SPOD [11], has gained attention not only because of the coherence of SPOD modes in both space and time, but also due to their relationship with resolvent analysis [12], where a linearization of the Navier-Stokes system is carried out and nonlinear terms are treated as an external forcing [13,14]. If such forcing is modeled as spatial white noise, optimal responses from resolvent analysis should match SPOD modes [15–17]. A similar correspondence with the analysis of the linearized Navier-Stokes operator also exists for POD modes, with the further assumption that nonlinearities are modeled as space-time white noise [18,19]. It is clear that such white-noise simplification is not realistic, and nonlinear terms have a “colored” structure, i.e., they display spatial coherence [20,21]. Nonetheless, the linearized operator often leads to gain separation between different amplification mechanisms, such that close agreement between leading SPOD and resolvent modes is obtained even if the forcing cannot be modeled as white noise. Detailed comparisons between SPOD modes and optimal responses from resolvent analysis have been carried out by Schmidt *et al.* [22] and Lesshafft *et al.* [23] for free jets and by Abreu *et al.* [24,25] for turbulent pipe and channel flow.

Such features explain the popularity of (S)POD to extract dominant coherent structures from databases of turbulent flows. However, although such modes are optimal in representing the turbulent kinetic energy, ultimately one often wishes to examine the effect of turbulent structures on other properties that are related but not directly defined by the kinetic energy. The aforementioned turbulent mixing, wall shear stress, and heat transfer are examples, as they depend on Reynolds stresses. These are products of different flow fluctuations, and are often studied at a given position in the flow using quadrant analysis or joint probability density functions [26], or in Fourier space using cospectra [27,28]. All these methods provide valuable information on Reynolds stresses, but, unlike a modal decomposition, results are obtained locally, for a given wall-normal position for instance. The product of different fluctuations cannot be expressed as an energy norm amenable to employing the proper orthogonal decomposition, and thus an appropriate modal decomposition is lacking. Besides standard Reynolds stresses, associated with the time-averaged Navier-Stokes equation, generalized Reynolds stresses for nonzero frequencies and wave numbers also arise in the study of more general nonlinear interactions exciting flow responses [14]. Due to their relevance in resolvent analysis [29,30], it would be important to optimally target combinations of flow fluctuations leading to high-amplitude nonlinear terms. A recent attempt in that direction is proposed by Schmidt [31], who decomposes the third-order flow statistics, represented by the bispectrum.

A method related to the present work is known as maximal covariance analysis, or MCA [32]. MCA seeks a pair of orthogonal bases to optimally represent the cross-covariance between two chosen quantities. For Reynolds stresses, the application of MCA would typically target the covariance between streamwise and wall-normal fluctuations, or between wall-normal velocity and temperature for thermal Reynolds stresses. This would lead to separate modes for each quantity involved in the covariance, with different temporal evolutions with some level of correlation. However, there may be physical arguments to infer that fluctuations in Reynolds stresses would be related to a single coherent structure. For instance, in wall-bounded flows near-wall fluctuations of streamwise and wall-normal velocity are related to streaks and quasistreamwise vortices, respectively [33]. These structures may be visualized separately, but they are dynamically linked by the equations of motion, such that streamwise vortices carry streamwise momentum leading to streaks via the lift-up mechanism [34,35]; consistently with such dynamical relation, application of POD leads to streamwise vortices and streaks appearing in a single coherent structure, or mode [8,25]. For thermal Reynolds stresses, one expects that the vertical velocity is related to temperature fluctuations, which drive plumes that emerge from the wall. The recent analysis by Krug *et al.* [28] shows that large-scale turbulent structures of vertical velocity and temperature in Rayleigh-Bénard convection

are strongly correlated, suggesting that thermal Reynolds stresses may be described by coherent modes involving velocity and temperature simultaneously.

It is the purpose of this work to develop an extension of POD, referred to as cross proper orthogonal decomposition (CPOD), that aims to optimally describe products of different flow quantities by extracting modes that maximize or minimize the spatial integral of such products. Thus, differently from MCA, we assume that Reynolds stresses may be described by a simultaneous temporal evolution of a pair of modes that comprise the two quantities involved in the covariance. As an example, we present CPOD of the Reynolds shear stress in a turbulent channel and examine which flow structures contribute to the overall Reynolds stress, and, more generally, to the mean velocity profile. The remainder of this paper is organized as follows. Section II formulates the problem and shows how CPOD modes and eigenvalues are obtained, showing the main properties of the modal decomposition. Section III exemplifies the procedure to decompose Reynolds stresses of turbulent channel flow, with a study of CPOD eigenvalues and eigenfunctions, and a comparison to other methods in the literature applied to the same database. Section IV completes the work with conclusions.

## II. DERIVATION OF CROSS PROPER ORTHOGONAL DECOMPOSITION

Let  $u(r)$  and  $v(r)$  be two zero-mean stochastic processes defined in a domain  $r \in \Omega$  such that each pair of realizations  $u_k$  and  $v_k$  are random variables, with the  $k$  subscript denoting the  $k$ th realization. We assume that each pair of realizations belongs to a Hilbert space  $\mathcal{H}(\Omega)$  with inner product

$$\langle u_k, v_k \rangle = \int_{\Omega} u_k v_k^* dr \approx \mathbf{v}_k^H \mathbf{W} \mathbf{u}_k, \quad (1)$$

where  $\mathbf{u}_k$  and  $\mathbf{v}_k$  are vectors whose entries are the discrete values of  $u_k(r)$  and  $v_k(r)$  at suitable  $n$  quadrature points, the superscripts  $*$  and  $H$  denote complex conjugate and Hermitian, respectively, and the weighting matrix  $W$  is a positive-definite Hermitian that represents quadrature weights. The underlying probability space is provided with an expectation operator  $\mathbb{E}[\cdot]$  that under the ergodicity hypothesis can be represented as a time average.

Moreover,  $\mathbf{u}_k$  and  $\mathbf{v}_k$  are considered to be related to the flow realizations  $\mathbf{x}_k$  and  $\mathbf{y}_k$  through observation matrices such that  $\mathbf{u}_k = \mathbf{C}_u \mathbf{x}_k$  and  $\mathbf{v}_k = \mathbf{C}_v \mathbf{y}_k$ . Choosing  $\mathbf{C}_u$  and  $\mathbf{C}_v$  as the identity matrix allows considering the full flow realizations. We would like to find a pair of deterministic functions  $(\psi, \theta)$  so that  $(\mathbf{C}_u \psi, \mathbf{C}_v \theta)$  is the best approximation of the expected relationship between  $u$  and  $v$ .

### A. Definitions and objectives

Consider two data matrices  $\tilde{\mathbf{X}} \in \mathbb{R}^{n \times N}$  and  $\tilde{\mathbf{Y}} \in \mathbb{R}^{n \times N}$  resulting from the collection of  $N$  flow realizations  $\mathbf{x}_i$  and  $\mathbf{y}_i$ . Each flow realization is a column vector with all flow variables at all spatial locations for a given time step, albeit  $\mathbf{x}$  and  $\mathbf{y}$  may eventually refer to different spatial wave numbers, for instance. The matrices  $\tilde{\mathbf{X}}$  and  $\tilde{\mathbf{Y}}$  are then given by

$$\tilde{\mathbf{X}} = [\mathbf{x}_1 \quad \mathbf{x}_2 \quad \cdots \quad \mathbf{x}_N] \quad (2a)$$

$$\tilde{\mathbf{Y}} = [\mathbf{y}_1 \quad \mathbf{y}_2 \quad \cdots \quad \mathbf{y}_N]. \quad (2b)$$

The means of snapshots  $\mathbf{x}_i$  and  $\mathbf{y}_i$  are respectively defined as

$$\bar{\mathbf{x}} = \mathbb{E}[\mathbf{x}] = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_i, \quad (3a)$$

$$\bar{\mathbf{y}} = \mathbb{E}[\mathbf{y}] = \frac{1}{N} \sum_{i=1}^N \mathbf{y}_i. \quad (3b)$$

The corresponding scaled perturbation matrices are

$$\mathbf{X} = \frac{1}{\sqrt{N-1}}(\tilde{\mathbf{X}} - \bar{\mathbf{x}}), \quad (4a)$$

$$\mathbf{Y} = \frac{1}{\sqrt{N-1}}(\tilde{\mathbf{Y}} - \bar{\mathbf{y}}), \quad (4b)$$

and the cross-covariance matrix between  $u$  and  $v$  can be finally written as

$$\mathbf{P}_{uv} = \mathbb{E}[\mathbf{C}_u \mathbf{x} \mathbf{y}^H \mathbf{C}_v^H] = \mathbf{C}_u \mathbf{X} \mathbf{Y}^H \mathbf{C}_v^H. \quad (5)$$

If  $u$  and  $v$  are two quantities involved in a component of Reynolds stresses, such as streamwise and wall-normal components, the main diagonal of the cross-covariance  $\mathbf{P}_{uv}$  contains such Reynolds stress. Although in some cases  $u$  and  $v$  may be complex-valued quantities, as when a preliminary Fourier decomposition is applied to the data, the primary interest is on the real part of the diagonal of  $\mathbf{P}_{uv}$ , since the Reynolds stress is a real quantity and the inverse Fourier transform naturally gives rise to the real part of  $\mathbf{P}_{uv}$ . This will be exemplified in Sec. III.

We label as  $\mathcal{I}$  the real part of spatial integral of the main diagonal of  $\mathbf{P}_{uv}$ . This may be obtained in discretized form using the quadrature weight matrix  $\mathbf{W}$  as

$$\mathcal{I} = \text{Re}[\text{Tr}(\mathbf{P}_{uv} \mathbf{W})] = \text{Re}[\text{Tr}(\mathbf{Y}^H \mathbf{C}_v^H \mathbf{W} \mathbf{C}_u \mathbf{X})], \quad (6)$$

where the cyclic property of the trace,  $\text{Tr}(\mathbf{A}_1 \mathbf{A}_2) = \text{Tr}(\mathbf{A}_2 \mathbf{A}_1)$ , is used. From this expression, the spatial integral  $\mathcal{I}$  may be obtained as the sum of inner products between corresponding realizations of  $u$  and  $v$ .

We here develop a procedure to obtain a modal decomposition of  $\mathbf{X}$  and  $\mathbf{Y}$  that is optimal in representing the spatial integral  $\mathcal{I}$ , or, equivalently, in representing the inner product between realizations of  $u$  and  $v$ . This is obtained in the next section.

## B. Optimization problem

The objective of CPOD is to extremize

$$\mathcal{J} = \langle \tilde{\mathbf{u}}, \tilde{\mathbf{v}} \rangle + \langle \tilde{\mathbf{v}}, \tilde{\mathbf{u}} \rangle = 2\text{Re}[\langle \tilde{\mathbf{u}}, \tilde{\mathbf{v}} \rangle], \quad (7)$$

where  $\tilde{\mathbf{u}}$  and  $\tilde{\mathbf{v}}$  are linear combinations of flow realizations  $\mathbf{x}_i$  and  $\mathbf{y}_i$  with different observation operators, such that

$$\tilde{\mathbf{u}} = \mathbf{C}_u \mathbf{X} \phi, \quad (8)$$

$$\tilde{\mathbf{v}} = \mathbf{C}_v \mathbf{Y} \phi, \quad (9)$$

with  $\phi$  as a column vector of  $N$  expansion coefficients. We constrain the maximization to consider  $\phi$  to be of unit Euclidean norm. Following Eq. (1), the inner product may be discretized as  $\langle \tilde{\mathbf{u}}, \tilde{\mathbf{v}} \rangle = \tilde{\mathbf{v}}^H \mathbf{W} \tilde{\mathbf{u}}$ .

In what follows we write a maximization problem for  $\mathcal{J}$ , but it will be seen that the same expression also leads to its minimization. The maximization problem becomes

$$\max_{\phi} \mathcal{J} = \max_{\phi} \frac{\phi^H (\mathbf{Y}^H \mathbf{C}_v^H \mathbf{W} \mathbf{C}_u \mathbf{X} + \mathbf{X}^H \mathbf{C}_u^H \mathbf{W} \mathbf{C}_v \mathbf{Y}) \phi}{\phi^H \phi}, \quad (10)$$

which, by defining

$$\mathbf{A} = \mathbf{Y}^H \mathbf{C}_v^H \mathbf{W} \mathbf{C}_u \mathbf{X} + \mathbf{X}^H \mathbf{C}_u^H \mathbf{W} \mathbf{C}_v \mathbf{Y}, \quad (11)$$

can be recognized as the maximization of the Rayleigh quotient

$$\max_{\phi} \mathcal{J} = \max_{\phi} \frac{\phi^H \mathbf{A} \phi}{\phi^H \phi}. \quad (12)$$

For a general matrix  $\mathbf{B}$ , its numerical range (or field of values)  $W(\mathbf{B})$  is defined as the subset of the complex plane

$$W(\mathbf{B}) = \left\{ \frac{\phi^H \mathbf{B} \phi}{\phi^H \phi} \mid \phi \in \mathbb{C}^n, \phi \neq 0 \right\}. \quad (13)$$

Because matrix  $\mathbf{A}$  is Hermitian, its numerical range is a line segment on the real axis delimited by its smallest and largest eigenvalues [36]. The cost function  $\mathcal{J}$  is, therefore, constrained by solutions of the eigenvalue problem

$$\mathbf{A}\phi_{(i)} = \lambda_{(i)}\phi_{(i)} \quad (14)$$

such that

$$\lambda_{\min} \leq \mathcal{J}(\phi) \leq \lambda_{\max}. \quad (15)$$

Since  $\mathbf{A}$  is Hermitian (although not positive definite in general), its eigenvalues are real and its  $N$  eigenfunctions  $\phi_{(i)}$ , when normalized, form an orthonormal basis satisfying  $\phi_{(j)}^H \phi_{(i)} = \delta_{ij}$ , where  $\delta_{ij}$  denotes the Kronecker delta (see finite-dimensional complex spectral theorem [37], for details).

Corresponding CPOD modes can be recovered as

$$\psi_{(i)} = \mathbf{X}\phi_{(i)}, \quad (16)$$

$$\theta_{(i)} = \mathbf{Y}\phi_{(i)}. \quad (17)$$

Unlike standard POD, in general, such CPOD modes  $\psi$  and  $\theta$  do not have an orthogonality relation; the orthogonality is restricted to the expansion coefficients  $\phi$ . For each pair of CPOD modes so defined, one has a contribution to  $\mathcal{J}$  given by  $\mathcal{J}_i = \lambda_i$ .

As CPOD modes are built from a linear combination of flow realizations, linear properties of realizations are maintained by the CPOD modes. For instance, for incompressible flows, CPOD modes are divergence-free. For the same reason, CPOD modes also satisfy linear homogeneous boundary conditions, such as zero velocity over walls. An extension to a frequency-domain version is straightforward, with realizations  $\mathbf{x}$  and  $\mathbf{y}$  taken from short-time Fourier transforms in the Welch method, as in SPOD [17]. Notice also that if  $\mathbf{X} = \mathbf{Y}$  and  $\mathbf{C}_u = \mathbf{C}_v$  the  $\mathbf{A}$  matrix is positive semidefinite, and the procedure above recovers POD modes. Hence, CPOD may be seen as a generalization of POD which allows modal expansions that target cross-covariances.

As a final remark, in the application studied in this work, where CPOD targets the (real-valued) Reynolds shear stress, the quantity of interest is the real part of the inner product between  $\tilde{\mathbf{u}}$  and  $\tilde{\mathbf{v}}$ , as will be exemplified in Sec. III. However, for generalized Reynolds stresses for a nonzero wave number or frequency, a complex-valued inner product may be of interest, as generalized Reynolds stresses are in general complex. In such situations, the complex-valued numerical range of Eq. (13) should be studied, using for instance the algorithm proposed by Schmidt [31].

### C. CPOD modes as a reconstruction basis

The nonorthogonality between modes  $\psi$  and  $\theta$  may be circumvented by the definition of an adjoint basis. For that matter, Eq. (14) may be manipulated as

$$\underbrace{\mathbf{X}\mathbf{X}^{-}}_{\tilde{\mathbf{A}}} \mathbf{X}\phi_{(i)} = \lambda_{(i)}\mathbf{X}\phi_{(i)}, \quad (18)$$

where  $\mathbf{X}^{-}$  is a pseudoinverse of  $\mathbf{X}$ . If  $\mathbf{X}$  has full rank, the pseudoinverse is a left inverse, satisfying  $\mathbf{X}^{-}\mathbf{X} = \mathbf{I}$ . The  $\psi$  modes are eigenfunctions of  $\tilde{\mathbf{A}} = \mathbf{X}\mathbf{X}^{-}$ , solutions of

$$\tilde{\mathbf{A}}\psi_{(i)} = \lambda_{(i)}\psi_{(i)}. \quad (19)$$

The above problem leads directly to the  $\psi$  modes, analogous to the ‘‘classical’’ POD method [as opposed to the method of snapshots that leads to Eq. (14)].

An adjoint eigenvalue problem may be obtained with  $\tilde{\mathbf{A}}^\dagger = \tilde{\mathbf{W}}^{-1} \mathbf{A}^H \tilde{\mathbf{W}}$ , with adjoint eigenvalues equal to the complex conjugate of the direct ones; since CPOD eigenvalues are real, the two problems have identical eigenvalues. The weight matrix  $\tilde{\mathbf{W}}$  is used to define an inner product in the space of realizations. The adjoint problem is

$$\tilde{\mathbf{A}}^\dagger \psi_{(i)}^\dagger = \lambda_{(i)} \psi_{(i)}^\dagger, \quad (20)$$

and the direct and adjoint modes satisfy a biorthogonality relation,

$$(\lambda_i - \lambda_j) (\psi_{(j)}^\dagger)^H \tilde{\mathbf{W}} \psi_{(i)} = 0, \quad (21)$$

which implies that pairs of direct and adjoint CPOD modes with different eigenvalue are orthogonal. This can be used to set up a biorthogonal projection, as discussed for instance in [38]. The same procedure may be applied to obtain a biorthogonality relation for the  $\theta$  modes, this time with  $\tilde{\mathbf{A}} = \mathbf{YAY}^-$ .

It is possible to obtain solutions of the CPOD direct and adjoint bases by this procedure, but this involves more significant computational costs, since for a flow with a large number of degrees of freedom, one needs to compute a large pseudoinverse in order to obtain  $\tilde{\mathbf{A}}$ , and the eigenvalue problem also has a larger dimension, given by the number of degrees of freedom of the flow [instead of the number of snapshots for Eq. (14), which is typically lower for flow databases]. This is not prohibitive, but requires careful choices of the numerical approach, such as the Arnoldi method for the eigenvalue problem [39].

Furthermore, consider the matrices

$$\Phi = [\phi_{(1)} \quad \phi_{(2)} \quad \cdots \quad \phi_{(N)}], \quad (22)$$

$$\mathbf{U} = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_N] = \mathbf{C}_u \mathbf{X} \Phi, \quad (23)$$

$$\mathbf{V} = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_N] = \mathbf{C}_v \mathbf{Y} \Phi. \quad (24)$$

Since  $\Phi$  is square and full rank, the multiplication by  $\Phi$  preserves the rank of the observed data matrices  $\mathbf{C}_u \mathbf{X}$  and  $\mathbf{C}_v \mathbf{Y}$ . Moreover,  $\Phi$  is a unitary matrix, such that  $\Phi^H \Phi = \Phi \Phi^H = \mathbf{I}$ , and the cross-covariance of Eq. (5) can be recovered as

$$\mathbf{P}_{uv} = \mathbf{C}_u \mathbf{X} \Phi \Phi^H \mathbf{Y}^H \mathbf{C}_v^H = \mathbf{C}_u \Psi \Theta^H \mathbf{C}_v^H, \quad (25)$$

where

$$\Psi = [\psi_{(1)} \quad \psi_{(2)} \quad \cdots \quad \psi_{(N)}], \quad (26)$$

$$\Theta = [\theta_{(1)} \quad \theta_{(2)} \quad \cdots \quad \theta_{(N)}]. \quad (27)$$

A rank- $M$  reconstruction of the cross-covariance  $\mathbf{P}_{uv}$  using CPOD modes amounts to Eq. (25) with the  $\Psi$  and  $\Theta$  matrices containing  $M$  modes each.

#### D. Synchronicity property

CPOD modes have, by construction, the property of being synchronous, i.e., corresponding  $\psi$  and  $\theta$  modes are expected to appear simultaneously in the data. In order to verify this claim, consider the problem of projecting the entries of the data matrix  $\mathbf{X}$  onto the CPOD modes  $\psi_{(i)}$ :

$$\Psi \gamma_i = \mathbf{x}_i, \quad (28a)$$

$$\gamma_i = (\Psi^\dagger)^H \tilde{\mathbf{W}} \mathbf{x}_i, \quad (28b)$$

where we have considered that direct and adjoint modes are normalized such that  $(\Psi^\dagger)^H \tilde{\mathbf{W}} \Psi = \mathbf{I}$ .

Grouping the results of performing such operation to all snapshots in a matrix  $\Gamma$  whose  $i$ th column is  $\gamma_i$ , and right multiplying by  $\Phi$  leads to

$$\Gamma \Phi = (\Psi^\dagger)^H \tilde{\mathbf{W}} \mathbf{X} \Phi = \mathbf{I}, \quad (29)$$

which leads to the conclusion that  $\Gamma = \Phi^{-1} = \Phi^H$ , since  $\Phi$  is a unitary matrix. An analogous algebraic manipulation would show that the projection coefficients of  $\mathbf{Y}$  onto  $\Theta$  are also given by  $\Phi^H$ .

The fact that the CPOD modes  $\psi_{(i)}$  and  $\theta_{(i)}$  have always the same coefficients in their respective data matrices indicates that the coherent structures they represent evolve synchronously in the data. High cross-covariances in the data are thus related to simultaneous high amplitudes of both  $\psi_{(i)}$  and  $\theta_{(i)}$  modes.

### E. Comparison with existing methods

Although the POD nomenclature is more common in areas related to fluid mechanics and mechanical engineering, the procedure is referred to differently in other fields. Examples include principal component analysis (PCA) in statistics and data sciences, Karhunen-Loève transform in signal processing, and empirical orthogonal functions in meteorology. The motivation behind all of these methods is to find an orthonormal basis that optimally describes the variance in a single set of data.

In this work, we deal with an alternative problem: how to optimally describe the cross-covariance matrix of two different sets of data. Several methods have been proposed in the literature to address this problem and the main distinction between them is their definition of optimality. In what follows, we discuss the differences between CPOD and some of these methods [5,32].

(1) *Canonical correlation analysis (CCA)*: CCA seeks pairs of vectors  $(\mathbf{p}, \mathbf{q})$  such that the expected correlation between  $\mathbf{p}^T \mathbf{x}$  and  $\mathbf{q}^T \mathbf{y}$  is successively maximized. It can be shown that  $\mathbf{p}$  and  $\mathbf{q}$  are related to right and left singular vectors of the cross-correlation matrix  $\mathbf{P}_{xx}^{-1/2} \mathbf{P}_{xy} \mathbf{P}_{yy}^{-1/2}$ .

(2) *Maximum covariance analysis (MCA)*: MCA seeks pairs of vectors  $(\mathbf{p}, \mathbf{q})$  such that the expected covariance between  $\mathbf{p}^T \mathbf{x}$  and  $\mathbf{q}^T \mathbf{y}$  is maximized. It can be shown that  $\mathbf{p}$  and  $\mathbf{q}$  are the right and left singular vectors of the cross-covariance matrix  $\mathbf{P}_{xy}$ .

(3) *Combined PCA (CPCA)*: CPCA concatenates  $\mathbf{x}$  and  $\mathbf{y}$  in the same vector  $\mathbf{z} = [\mathbf{x} \ \mathbf{y}]^T$  and applies the PCA methodology to  $\mathbf{z}$ . While CPOD maximizes  $\text{Re}\{\text{trace}(\mathbf{P}_{xy})\}$  contained in a subspace of unitary dimension (i.e., spanned by the leading CPOD mode), CPCA maximizes  $[\text{trace}(\mathbf{P}_{xx}) + \text{trace}(\mathbf{P}_{yy})]$ .

(4) *Redundancy analysis (RA)* [40]: RA seeks a vector  $\mathbf{p}$  such that the expected correlation between  $\mathbf{p}^T \mathbf{x}$ , regarded as a linear function of the predictor variables, and  $\mathbf{y}$ , regarded as the response variables, is successively maximized. This optimality criterion is equivalent to solving the generalized eigenvalue problem:

$$\mathbf{R}_{xy} \mathbf{R}_{yx} \mathbf{p} = \sigma \mathbf{R}_{xx} \mathbf{p}, \quad (30)$$

where  $\mathbf{R}_{xy}$  is the correlation matrix between variables  $\mathbf{x}$  and  $\mathbf{y}$  and  $\sigma$  is the generalized eigenvalue. Swapping  $x$  with  $y$  in Eq. (30), one can then obtain the vector  $\mathbf{q}$  associated to the data  $\mathbf{y}$ . Thacker [41] proposed a RA variant termed *principal predictor* in which the vector  $\mathbf{p}$  is given by the solutions to the generalized eigenvalue problem:

$$\mathbf{P}_{xy} [\text{diag}(\mathbf{P}_{yy})]^{-1} \mathbf{P}_{yx} \mathbf{p} = \sigma \mathbf{P}_{xx} \mathbf{p}. \quad (31)$$

(5) *Observable inferred decomposition (OID)* [42]: OID seeks a modal representation of the state  $\mathbf{x}$  that optimizes the representation of the variance of an observable  $\mathbf{y} = \mathbf{C}\mathbf{x}$ . It can be considered a generalization of the POD methodology in which the optimization is performed using a weighted norm, in which the weight matrix is based on the linear stochastic estimation between the state and the observable.

Among the methods above, CCA and MCA would appear to be good alternatives to CPOD. However, they lack the synchronicity property presented in Sec. IID. As a consequence, there is no guarantee that optimal CCA and MCA modes can be found simultaneously in the data. Correlations or covariances may be maximized by CCA or MCA modes, which serve as optimal bases for the *statistics*, but such modes cannot be associated to instantaneous coherent structures in a straightforward manner. For further discussion, CPOD results will be compared to MCA in Sec. III F using a snapshot method for MCA as derived in the Appendix.

### III. EXAMPLE: APPLICATION FOR REYNOLDS STRESSES OF TURBULENT CHANNEL FLOW

#### A. Formulation of the problem

We would like to find modes  $\phi$ , or equivalently  $\psi$ , that optimally represent the Reynolds shear stress  $\tau_{xy} = \overline{u'v'}$  in a turbulent channel flow database, where  $u'$  and  $v'$  denote, respectively, velocity fluctuations in streamwise ( $x$ ) and wall-normal ( $y$ ) directions and the overline denotes averaging in time and in wall-parallel directions  $x$  and  $z$ . Consider a problem with periodic boundary conditions in  $x$  and  $z$ . In this situation, a Fourier decomposition in space is carried out for discrete wave numbers, and Fourier series coefficients are obtained as

$$\hat{u}(\alpha, y, \beta, t) = \frac{1}{L_x L_z} \int_0^{L_z} \int_0^{L_x} u(x, y, z, t) e^{-i(\alpha x + \beta z)} dx dz, \quad (32a)$$

$$\hat{v}(\alpha, y, \beta, t) = \frac{1}{L_x L_z} \int_0^{L_z} \int_0^{L_x} v(x, y, z, t) e^{-i(\alpha x + \beta z)} dx dz. \quad (32b)$$

For the channel geometry, only the Reynolds shear stress  $\overline{u'v'}$ ( $y$ ) is related to the mean velocity profile. The averaging in  $x$  and  $z$  implies that one is interested in the  $(\alpha, \beta) = (0, 0)$  component of Reynolds stress. Thus, using the convolution theorem, we can represent the averaged Reynolds stress as a sum of contributions of  $(-\alpha, -\beta)$  and  $(\alpha, \beta)$  wave numbers, whose sum leads to the streamwise- and spanwise-averaged Reynolds stress. Such sum is given by

$$\overline{u'v'}(y) = \sum_{\alpha=-\alpha_N}^{\alpha_N} \sum_{\beta=-\beta_N}^{\beta_N} \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \hat{u}(-\alpha, y, -\beta, t) \hat{v}(\alpha, y, \beta, t) dt, \quad (33)$$

where  $\alpha_N$  and  $\beta_N$  denote the highest considered wave numbers (typically the Nyquist values for data sampled in a uniform grid). Since  $u'$  and  $v'$  are real variables, from the reality condition,

$$\hat{u}(\alpha, y, \beta, t) = \hat{u}^H(-\alpha, y, -\beta, t), \quad (34)$$

$$\hat{v}(\alpha, y, \beta, t) = \hat{v}^H(-\alpha, y, -\beta, t), \quad (35)$$

which allows rewriting Eq. (33) as a sum restricted to positive  $\beta$ , as

$$\begin{aligned} \overline{u'v'}(y) &= 2\text{Re} \left[ \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \sum_{\alpha=-\alpha_N}^{\alpha_N} \sum_{\beta=\beta_1}^{\beta_N} \hat{u}(\alpha, y, \beta, t) \hat{v}^H(\alpha, y, \beta, t) dt \right] \\ &\quad + \text{Re} \left[ \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \sum_{\alpha=-\alpha_N}^{\alpha_N} \hat{u}(\alpha, y, \beta = 0, t) \hat{v}^H(\alpha, y, \beta = 0, t) dt \right] \\ &= \sum_{p,q} \mathbb{E} \{ 2\text{Re} [ \hat{u}_r(\alpha_p, y, \beta_q) \hat{v}_r^H(\alpha_p, y, \beta_q) ] \}, \end{aligned} \quad (36)$$

with  $\beta_1$  as the first positive spanwise wave number. Equation (36) expresses the Reynolds stress as a sum of expected values for each wave-number combination  $(\alpha_p, \beta_q)$ ; for  $\beta = 0$ ,  $\hat{u}_r$  and  $\hat{v}_r$  should be divided by  $\sqrt{2}$  to comply with Eq. (36). The quantity  $\mathbb{E} \{ 2\text{Re} [ \hat{u}_r(\alpha_p, y, \beta_q) \hat{v}_r^H(\alpha_p, y, \beta_q) ] \}$  is the

contribution of wave numbers  $(\alpha_p, \beta_q)$  to the total Reynolds stress, and each wave-number pair can be separately analyzed using CPOD, as follows. A spatial integration over  $y$  for a wave-number pair leads to a functional

$$\int \overline{u'v'} dy = \sum_{p,q} \mathcal{F}_{p,q} = \sum_{p,q} \mathbb{E} \left\{ 2\text{Re} \left[ \int \hat{u}_r(\alpha_p, y, \beta_q) \hat{v}_r^H(\alpha_p, y, \beta_q) dy \right] \right\}, \quad (37)$$

where the contribution of each wave-number pair  $\alpha_p$  and  $\beta_q$ , written as

$$\mathcal{F}_{p,q} = \mathbb{E}[2\text{Re}(\hat{u}_r(\alpha_p, y, \beta_q), \hat{v}_r(\alpha_p, y, \beta_q))], \quad (38)$$

is related to an inner product along the lines of Eq. (7), which may be independently studied using CPOD.

### B. Database and conventions

We consider turbulent channel flow at friction Reynolds number  $\text{Re}_\tau = 179$  to exemplify an application of CPOD. A direct numerical simulation (DNS) was performed, using the pseudospectral code CHANNELFLOW [43]. Periodic boundary conditions are used for streamwise ( $x$ ) and spanwise ( $z$ ) directions, with no-slip conditions on the channel walls. Discretization in the  $y$  direction is based on Chebyshev polynomials. The computational domain is  $(4\pi, 2, 2\pi)$  times the half channel height in streamwise, wall-normal, and spanwise directions, respectively. Further information and validation results are presented in [44].

For turbulent channel flow, two normalizations can be used. The first is based on the outer quantities, bulk velocity and the half channel height, whereas a second normalization can be defined based on an inner, viscous scaling based on the friction velocity  $u_\tau$  and kinematic viscosity. Inner-scaled quantities are denoted by a + superscript; otherwise, outer scaling is implied.

For a given pair of streamwise and spanwise wave numbers  $\alpha$  and  $\beta$ , flow realizations were taken for the velocity fluctuations, following

$$\mathbf{x}_i = \mathbf{y}_i = \begin{bmatrix} \hat{\mathbf{u}}^+(\alpha, y, \beta, t_i) \\ \hat{\mathbf{v}}^+(\alpha, y, \beta, t_i) \\ \hat{\mathbf{w}}^+(\alpha, y, \beta, t_i) \end{bmatrix}, \quad (39)$$

where vectors  $\hat{\mathbf{u}}^+$ ,  $\hat{\mathbf{v}}^+$ , and  $\hat{\mathbf{w}}^+$  contain inner-scaled, velocity fluctuations in streamwise, wall-normal, and spanwise directions, respectively, for a given wave number. Observation operators  $\mathbf{C}_u = [\mathbf{I} \ \mathbf{0} \ \mathbf{0}]$  and  $\mathbf{C}_v = [\mathbf{0} \ \mathbf{I} \ \mathbf{0}]$  were used to extract streamwise and wall-normal velocity components, respectively, to construct the inner product leading to the Reynolds stress. Consideration of the three velocity components in the matrix of realizations leads to modes with the three velocity components satisfying the incompressibility condition. A total of 250 realizations were taken, spanning 500 outer convective time units. The reduction of this number to half led to nearly identical results.

Turbulent channel flow is symmetric in  $y$ , and, similarly to POD modes, CPOD modes should be split into even and odd functions. Such decomposition is performed *a priori*, with even and odd modes defined based on  $u$  and  $w$ . The wall-normal fluctuations  $v$  have the opposite parity of  $u$  and  $w$ . To apply CPOD, a weight matrix  $\mathbf{W}$  was defined using Clenshaw-Curtis quadrature [45] using outer coordinates. Since modes are symmetric or antisymmetric by construction, only half of the domain was considered in  $\mathbf{W}$ . Moreover, the Reynolds shear stress is antisymmetric in channel flow, and thus the  $y$ -integrated Reynolds stress is zero; thus, spatial integration over only half a channel is appropriate in the present case.

CPOD was carried out for the first 17 streamwise wave numbers  $\alpha$  and 32 spanwise wave numbers  $\beta$  that are allowed by the periodic computational box. To span the whole Fourier domain, the computation was also carried out for 16 negative values of  $\alpha$ . Since turbulent channel flow has no preferred spanwise sense for fluctuations, eigenvalues should satisfy  $\sigma(\alpha, \beta) = \sigma(-\alpha, \beta)$  if a

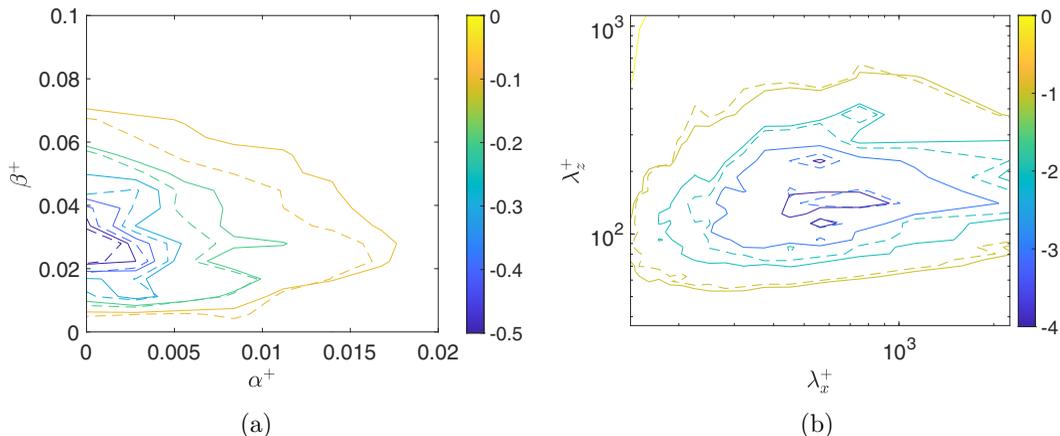


FIG. 1. Wave-number spectra. Full lines show the Reynolds stress contribution of each wave number, and dashed lines show the contribution of the leading CPOD eigenvalue. Results shown for the odd-mode contribution. (a) Reynolds stress and (b) premultiplied Reynolds stress.

sufficient number of snapshots is taken for converged statistics. Since CPOD eigenvalues may be positive or negative, they were ranked by their absolute value in descending order.

### C. Wave-number spectra of Reynolds stresses

The contributions of wave numbers to the overall Reynolds stress can be assessed by examining the CPOD eigenvalues as a function of wave number. Figure 1 shows both leading eigenvalue  $\sigma_1$  and the sum of eigenvalues  $\sum_i \sigma_i$  as a function of inner-scaled wave numbers  $\alpha^+$  and  $\beta^+$ . It also shows such spectra in premultiplied form [46], as a function of streamwise and spanwise wavelengths  $\lambda_x^+ = 2\pi/\alpha^+$  and  $\lambda_z^+ = 2\pi/\beta^+$ . The first relevant feature of these plots is that the leading eigenvalue  $\sigma_1$  (dashed contour lines) is already close to the full sum (full lines), which shows that the CPOD spectrum is dominated by the first mode. The nonpremultiplied spectrum has a peak for  $(\alpha^+, \beta^+) = (0, 0.028)$ , corresponding to  $(\lambda_x^+, \lambda_z^+) = (\infty, 225)$ . These are structures that span the whole computational domain in the streamwise direction. In the premultiplied representation, the contribution of such structures vanishes, as it is multiplied by  $\alpha^+ \beta^+$ , and the peak contribution is obtained for  $(\lambda_x^+, \lambda_z^+) = (750, 140)$ . Both wave-number combinations will be further examined in the following sections.

### D. Low-rank reconstructions of Reynolds stress and mean field

We next examine how a superposition of modes is able to reconstruct the overall Reynolds stress in the DNS. For that matter, Eq. (36) is used considering a low-rank reconstruction of the Reynolds stress. Figure 2(a) shows the total Reynolds stress from the DNS with a black line. For reference, the total stress is shown with a blue dashed line. At  $y = -0.85$ , corresponding to a wall-normal distance of  $y^+ \approx 27$ , the Reynolds stress already accounts for almost all the total stress. The reconstructions with a small number of CPOD modes are shown in the same figure, with  $N = 1, 2, \dots, 5$  modes, ordered based on the absolute value of the eigenvalue, considered for each wave number in the database. For each case, both even and odd modes are considered in the reconstruction, and  $N = 1$  implies that the first even and the first odd mode were considered for each wave number in the database, with contributions of all wave numbers superposed to obtain the overall Reynolds shear stress. As hinted by Fig. 1, the first CPOD mode pair of each wave number already leads to a close reconstruction of the overall Reynolds stress, recovering the peak value. The addition of subsequent modes leads to nonmonotonic convergence toward the DNS values. This can be understood by

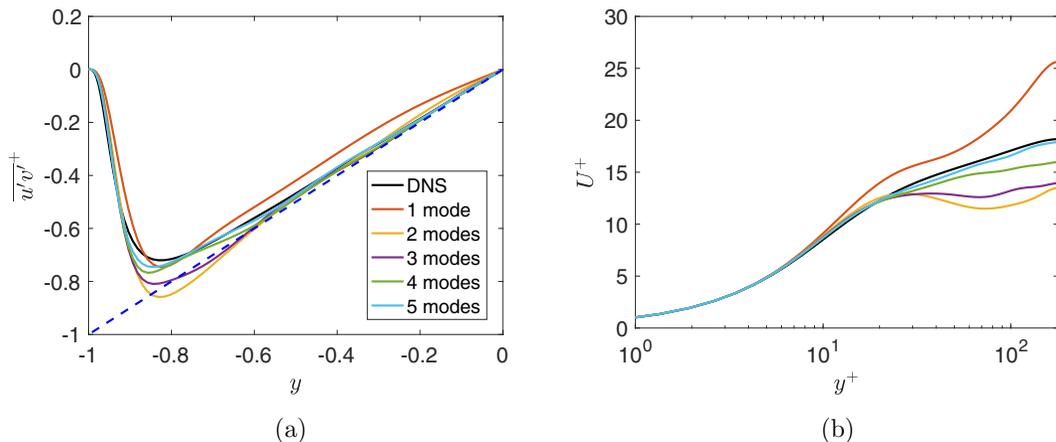


FIG. 2. Reconstruction of (a) Reynolds stress and (b) mean flow using a variable number of CPOD modes for each wave number of the database, compared to reference DNS results. The blue dashed line in (a) shows the total stress (equal to  $y$ ) for reference.

recalling that CPOD eigenvalues may be positive or negative. Negative (positive) eigenvalues lead to negative (positive) contributions to  $\overline{u'v'}$ . From the observed convergence, it is inferred that the two leading CPOD modes have an overall negative contribution to the Reynolds stress, whereas the third and fourth modes have a mostly positive contribution, leading to a less negative reconstructed peak. Superposition of five pairs of modes is shown to lead to a close reconstruction of the DNS stresses.

The reconstructed Reynolds stress allows a computation of the reconstructed mean flow  $U^+$  by

$$\frac{dU^+}{dy^+} = 1 - \frac{y^+}{\text{Re}_\tau} + \overline{u'v'}(y^+), \quad (40)$$

which can be integrated from  $y^+ = 0$  in a straightforward manner. Figure 2(b) shows reconstructions of the mean flow using  $\overline{u'v'}$  approximated by a small number of CPOD modes. Although a single mode pair already leads to a reasonable reconstruction of  $\overline{u'v'}$ , the corresponding mean-flow reconstruction is shown to suffer from Reynolds stresses that are below the DNS values. The inclusion of higher-order modes improves immediately the mean-flow representation up to  $y^+ = 30$ , and with five mode pairs the DNS profile is closely reconstructed.

### E. Eigenvalues and eigenfunctions for peak wave numbers

We now turn our attention to the peak wave numbers of Fig. 1, which are further analyzed by looking at leading CPOD eigenvalues and eigenfunctions. In the remainder of this paper, only the odd modes are studied; even modes display similar behavior and are not shown. Figure 3 shows leading eigenvalues for  $(\lambda_x^+, \lambda_z^+) = (\infty, 225)$  and  $(750, 140)$ , ranked by absolute value. The leading mode for both cases has a negative eigenvalue, consistent with the negative Reynolds stress at the lower half channel. However, starting from the third or fourth mode, positive eigenvalues are obtained, albeit with comparatively low absolute value, consistent with the oscillatory convergence seen in Fig. 2.

Corresponding eigenfunctions are shown in Figs. 4 and 5, which refer to the peaks of the spectra without and with premultiplication, respectively. The figures show the leading two modes and the first subsequent mode with positive eigenvalue. The two leading modes are similar to overall features seen in POD [8] and SPOD [25] of wall-bounded flows. Since both cases correspond to  $\lambda_x^+ \ll \lambda_z^+$ , the structures in physical space are elongated in the streamwise direction. Velocity fluctuations are

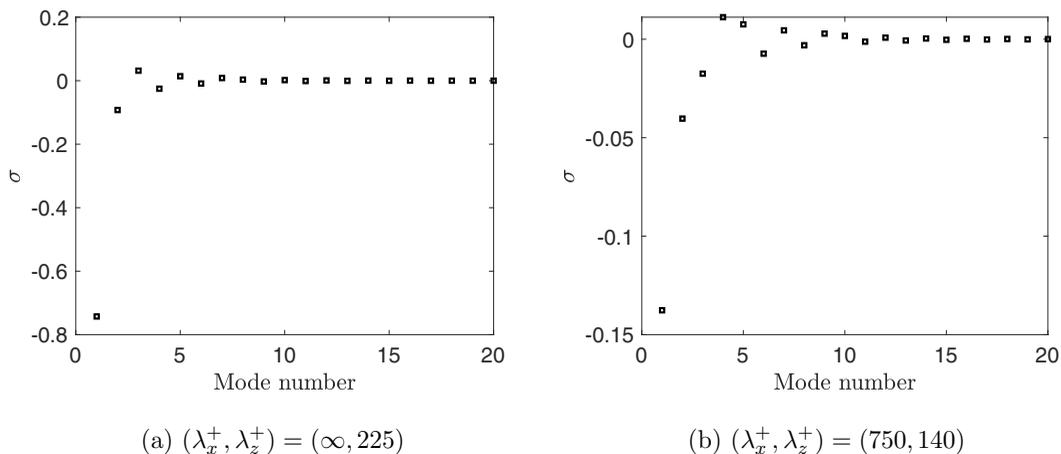


FIG. 3. Eigenvalues for wave numbers corresponding to (a) the peak of Reynolds stress and (b) the premultiplied peak of Reynolds stress.

dominated by streaks of streamwise velocity  $u'$ , and  $v'$  and  $w'$  components represent streamwise vortices, with  $v'$  in phase opposition with  $u'$ . This is consistent with the lift-up mechanism, as positive wall-normal fluctuations close to the lower wall ( $v' > 0$ ) carry low-speed fluid to high-speed regions, leading to negative fluctuations in  $u$ , in a process referred to as ejections, corresponding to Q2 ( $u' < 0, v' > 0$ ) in quadrant analysis [26]. Due to the spanwise periodicity of the mode, the opposite happens half a wavelength away, in a sweep corresponding to quadrant Q4 ( $u' > 0, v' < 0$ ). For the second mode, a pair of structures with phase opposition in  $y$  is observed, similarly to the

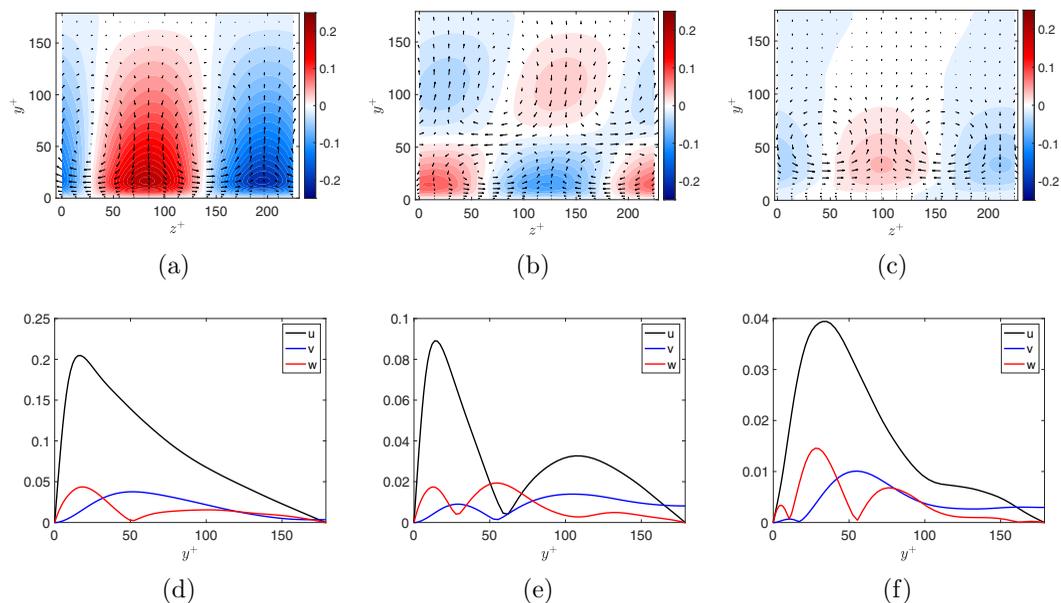


FIG. 4. Eigenvalues and sample eigenfunctions for  $(\lambda_x^+, \lambda_z^+) = (\infty, 225)$ , the peak of the non-premultiplied spectrum of Reynolds stress. Colors:  $u$ ; arrows:  $v$  and  $w$ . (a) Mode 1, (b) Mode 2, (c) Mode 3, (d) Mode 1 amplitudes, (e) Mode 2 amplitudes, and (f) Mode 3 amplitudes.

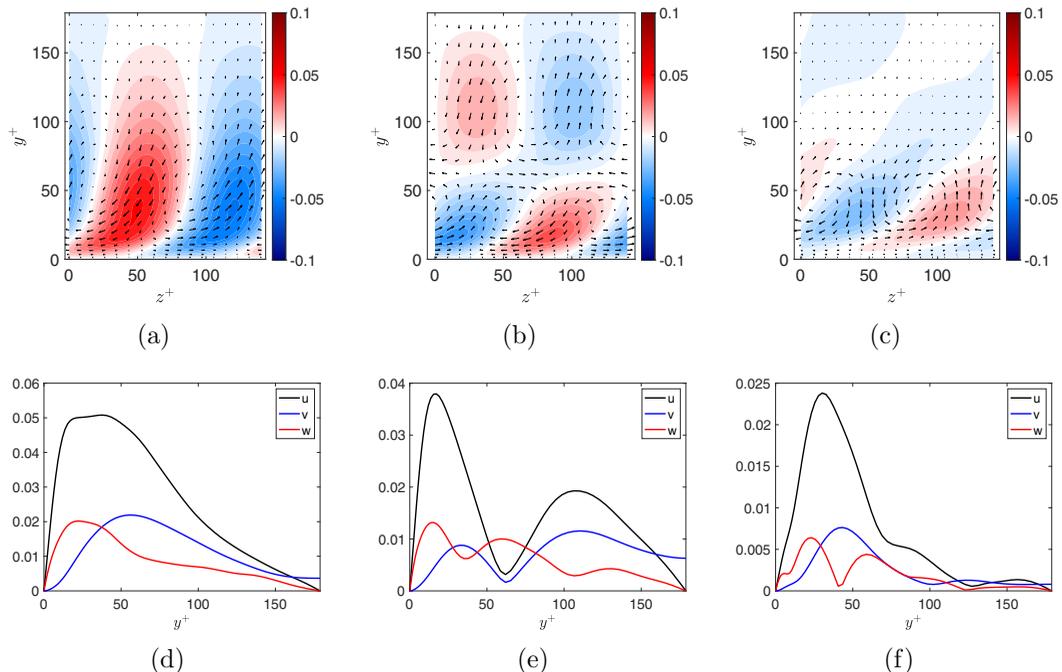


FIG. 5. Sample eigenfunctions for  $(\lambda_x^+, \lambda_z^+) = (750, 140)$ , the peak of the premultiplied spectrum of Reynolds stress. Colors:  $u$ ; arrows:  $v$  and  $w$ . (a) Mode 1, (b) Mode 2, (c) Mode 4, (d) Mode 1 amplitudes, (e) Mode 2 amplitudes, and (f) Mode 4 amplitudes.

observations of (S)POD modes by [8] and [25], again with  $u'$  and  $v'$  in phase opposition. Both eigenfunctions have thus a negative contribution to the Reynolds stress  $\overline{u'v'}$ .

The third mode for  $(\lambda_x^+, \lambda_z^+) = (\infty, 225)$  and the fourth mode from  $(\lambda_x^+, \lambda_z^+) = (750, 140)$  depart from such behavior. Streaks and streamwise vortices are again observed, but this time  $u$  and  $v$  are in phase. This shows that in some realizations of the velocity field streamwise vortices and streaks are in a configuration that is opposite to the standard lift-up scenario. This is not a dominant phenomenon, as it appears only on higher-order modes. Such modes correspond to Q1 and Q3 in quadrant analysis, which, despite their lower magnitude, have relevant contributions to the Reynolds stress [26].

The observation of these sample modes explains the oscillatory convergence of Reynolds stress seen in Fig. 2. CPOD sheds light on turbulent structures with contributions to the averaged Reynolds stress, and it is seen that although most structures have negative contributions to  $\overline{u'v'}$  for the lower half channel, some structures have positive contributions, and the overall Reynolds stress results from constructive and destructive interferences among the various modes. As shown in Sec. III D, such higher-order modes are relevant to obtain the mean-flow profile.

### F. Comparison with other decompositions

We now compare the CPOD modes studied in the previous section with two reference modal decompositions of the literature. As in the previous section, we restrict the comparison to odd modes, as even modes display a similar behavior. CPOD is compared to standard POD, targeting the turbulent kinetic energy, and to MCA, targeting the full covariance  $\mathbf{P}_{uv}$  in Eq. (5). Notice that MCA leads to pairs of left and right eigenfunctions with maximal contribution to the covariance, but without the synchronicity property of CPOD, as discussed in Sec. II D. MCA modes were obtained

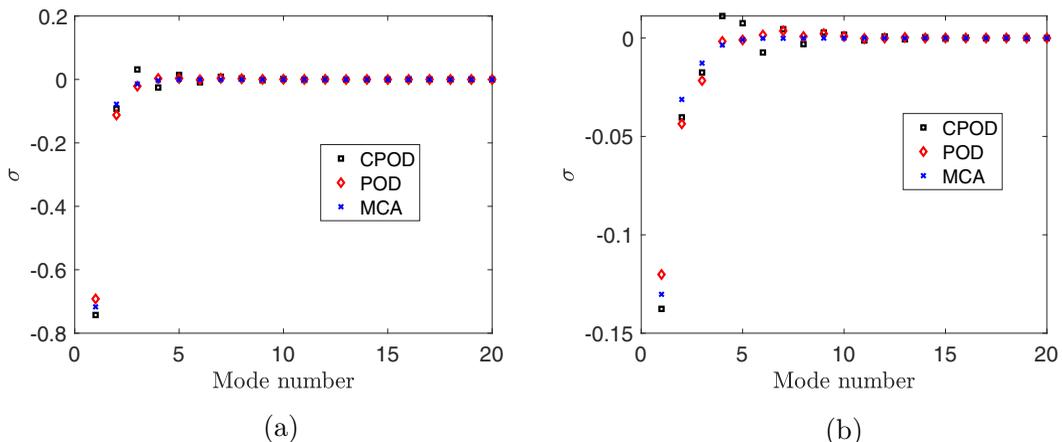


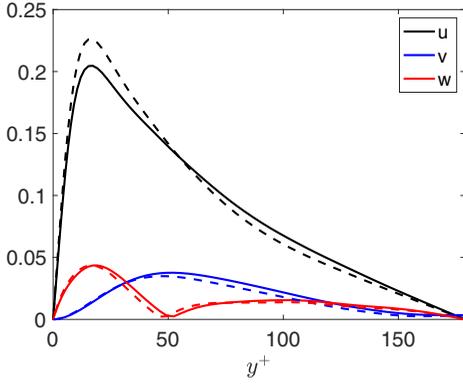
FIG. 6. Contribution of each mode to spatial integral of Reynolds stress. (a) Peak of non-premultiplied spectrum and (b) peak of premultiplied spectrum.

with a snapshot method, derived in the Appendix. The comparison between methods focuses on the wave numbers of peak Reynolds stress, with and without premultiplication, studied in Sec. III E.

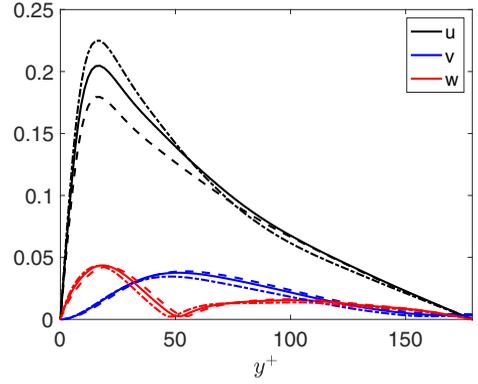
As each technique targets a different quantity, eigenvalues are not directly comparable. We calculate the real part of the spatial integral of Reynolds shear stress resulting from each modal decomposition, referred to as  $\sigma_i$  in analogy with the CPOD eigenvalue. The results are shown in Fig. 6. For both wave numbers, we notice that the integrated Reynolds stress of the leading CPOD mode is larger than both the POD eigenfunction and the MCA mode pair. Regarding POD, this is expected, as the optimality of CPOD ensures that no other spatial function has a larger contribution to the integrated Reynolds stress, as derived in Sec. II B. However, this may not be the case when compared to the MCA results, since MCA relaxes the requirement of synchronicity and allows separate time coefficients for the left and right eigenfunctions. The single leading CPOD nonetheless leads to a higher contribution to the integrated Reynolds stress than the leading MCA mode pair. This is likely due to the consideration of the real part of the Reynolds stresses in CPOD, and not in MCA. As only the real part of the Reynolds shear stress contributes to the overall value, as developed in Sec. III, this provides an advantage for CPOD over MCA, with slightly larger Reynolds stresses for the leading mode in Fig. 6.

A comparison of leading eigenfunctions is shown in Fig. 7. POD modes naturally involve the three velocity components. As discussed in the Appendix, MCA modes would in principle comprise solely the quantities in the cross-covariance of interest, and left and right MCA modes would only be formed by  $u$  and  $v$  components, respectively. However, by computing MCA modes with the snapshot method described in the Appendix, MCA modes with all degrees of freedom, and hence all velocity components, may be computed, for a more complete comparison with CPOD and POD modes. The leading eigenfunctions of each method are thus compared in Fig. 7 in all their velocity components.

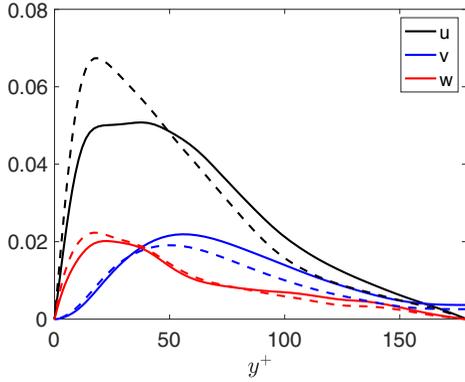
The results of Fig. 7 show modes that are overall similar for all methods. In all cases, modes display higher amplitudes of  $u$ , representing streaks, and lower-amplitude cross-flow velocities  $v$  and  $w$  representing nearly streamwise vortices. The appearance of such structures in all methods is expected, as these are the dominant structures for the flow at hand [24,33]. However, the details of the eigenfunctions are related to the optimal properties of each method. POD modes target the turbulent kinetic energy (TKE), and the leading POD mode has a streak and a streamwise vortex that optimally represent the TKE from the data. This is contrasted with the leading CPOD mode, which optimally represents the Reynolds shear stress. In the first CPOD mode shown in Figs. 7(a) and 7(c), we observe a streak with lower peak amplitude but more distributed in  $y$  if compared to



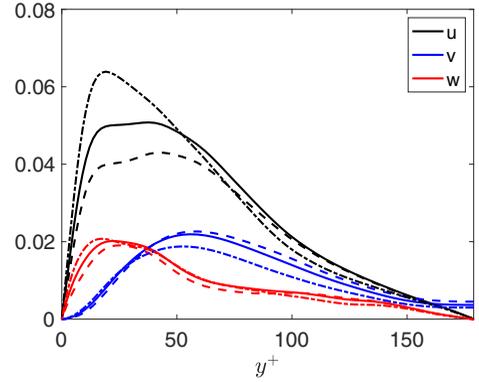
(a) CPOD and POD modes, peak of non-premultiplied spectrum



(b) CPOD and MCA modes, peak of non-premultiplied spectrum



(c) CPOD and POD modes, peak of premultiplied spectrum



(d) CPOD and MCA modes, peak of premultiplied spectrum

FIG. 7. Comparison between CPOD modes (full lines in all subfigures), POD modes [dashed lines in (a) and (c)], and MCA modes (dashed and dash-dotted lines in (b) and (d)]. Dashed and dash-dotted lines in (b) and (d) refer respectively to left and right MCA modes.

the POD mode, particularly for the peak of premultiplied Reynolds stress, shown in Fig. 7(c). For higher  $y^+$ , the CPOD mode also has a higher  $v$  component, related to a stronger streamwise vortex. These two properties lead to a higher contribution to the Reynolds stress, as seen in Fig. 6, an effect that is more marked for the peak of the premultiplied spectrum.

A reconstruction of the overall Reynolds stress using POD modes is shown in Fig. 8(a). As in Fig. 2, we have taken a number of one to five pairs of POD modes, one even and one odd, for all wave numbers in the database. The results show the convergence to the overall Reynolds stress as more modes are superposed. POD modes eventually lead to a reconstruction of the Reynolds stress, but at a rate slower than the one observed for CPOD. This can be seen by comparing the POD-based reconstruction with the one based on CPOD (dashed lines in Fig. 8, which are taken from Fig. 2). For a given number of modes, a reconstruction using CPOD leads to results that are closer to the overall Reynolds stress. This is also reflected in the mean-flow reconstructions shown in Fig. 8(b):

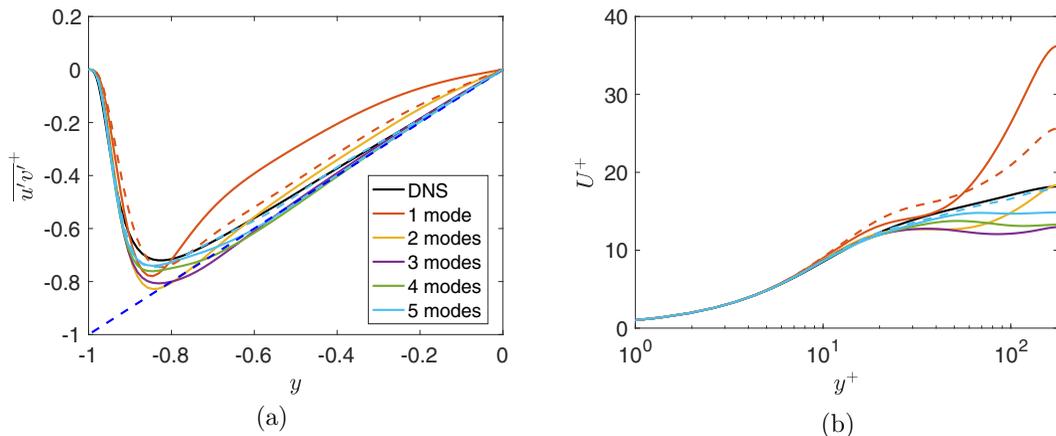


FIG. 8. Reconstruction of (a) Reynolds stress and (b) mean flow using a variable number of POD modes for each wave number of the database, compared to reference DNS results. The blue dashed line in (a) shows the total stress (equal to  $y$ ) for reference. Other dashed lines show one- and five-mode reconstructions using CPOD, taken from Fig. 2, for comparison.

the faster convergence of Reynolds stress with CPOD leads to a better recovery of the mean velocity profile, as seen in the one- and five-mode reconstructions shown in the figure.

A comparison with MCA modes, shown in Fig. 7, has further subtlety. The Reynolds stress in MCA is obtained by taking  $u$  from the left eigenfunction, and  $v$  from the right eigenfunction. However, both left and right eigenfunctions comprise streaks and streamwise vortices, of similar shape to CPOD and POD modes. MCA modes, shown in Fig. 7, display streaks in the left eigenfunction that are more spread in  $y$ , but streamwise vortices in the right eigenfunction with a lower  $v$  component. MCA modes then turn out to be optimal in representing the entire  $u-v$  cross-covariance, but are actually suboptimal in representing the real part of the Reynolds shear stress, which is the quantity of interest to determine the mean velocity profile.

A further complication is that, even though the target is the cross-covariance between  $u$  and  $v$ , MCA eigenfunctions obtained using the snapshot method of the Appendix have the three velocity components, and both left and right MCA modes comprise slightly different streaks and streamwise vortices. One could possibly expect the left and right eigenfunctions to be streak and vortex dominated, respectively, but this is not the case, and left and right modes involve streaks and streamwise vortices at the same time. As discussed in the Introduction, a split of wall turbulence into separate streaks and vortices may actually be artificial, as the three velocity components are intrinsically linked by the continuity and momentum equations. For the present application, MCA leads to a pair of quite similar, but asynchronous structures, whose combination approximates the cross-covariance  $\mathbf{P}_{uv}$  and hence the Reynolds stress.

Moreover, in an analysis based on MCA one should keep in mind that two different structures are obtained in order to represent a cross-covariance of interest. This leads to separate modes that may be used in the reconstruction of the data matrices  $\mathbf{X}$  and  $\mathbf{Y}$ , as detailed in the Appendix. In the present case, the Reynolds shear stress is obtained from data matrices  $\mathbf{X} = \mathbf{Y}$ ; hence, an expansion of the Reynolds stress would require projections of the same data matrix  $\mathbf{X}$  onto two MCA bases  $\mathbf{U}$  and  $\mathbf{V}$ . Individual left or right eigenfunctions are divergence-free, as shown in the Appendix; however, if  $u$  and  $v$  are taken from the left and right eigenfunctions, respectively, the resulting field is no longer divergence-free. Based on these points one cannot think of a superposition of MCA modes reconstructing the velocity field and hence the Reynolds stress, as in POD and CPOD; in this case, it is more appropriate to think of a reconstruction of the cross-covariance  $\mathbf{P}_{uv}$  using the two bases,

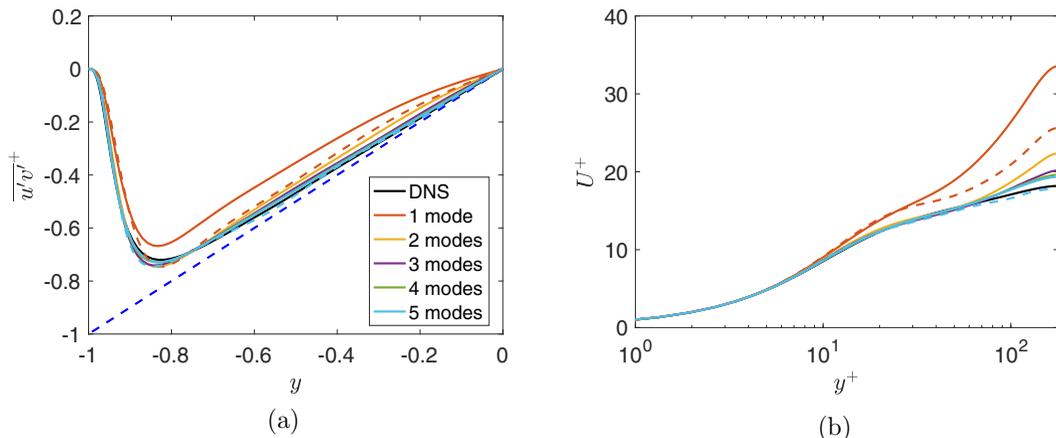


FIG. 9. Reconstruction of (a) Reynolds stress and (b) mean flow using a variable number of MCA mode pairs for each wave number of the database, compared to reference DNS results. The blue dashed line in (a) shows the total stress (equal to  $y$ ) for reference. Other dashed lines show one- and five-mode reconstructions using CPOD, taken from Fig. 2, for comparison.

which lead to dual reconstructions of the turbulent field. The interpretation is further complicated by the aforementioned similarity, but lack of synchronicity, between left and right MCA modes.

These aspects highlight CPOD as a pertinent approach to optimally decompose cross-covariances onto synchronous modes, which, for the Reynolds shear stress studied here (and for whenever  $\mathbf{X} = \mathbf{Y}$ ), leads to a single structure. For the turbulent channel at hand such structure comprises a streamwise vortex and a streak, which are intrinsically coupled. CPOD is thus a more compact representation, which circumvents the difficulties in the interpretation of MCA results for the present case. Generalized Reynolds stresses built from data matrices  $\mathbf{X} \neq \mathbf{Y}$  may lead to a more straightforward application of MCA.

Keeping in mind the aforementioned caveats in the interpretation of MCA results, Fig. 9 presents a reconstruction of Reynolds shear stress and mean flow using MCA mode pairs, with CPOD results shown with dashed lines for reference. CPOD has better reconstructions if a single mode per wave number is considered, as it targets the real part of the cross-covariance. The five-mode reconstruction is also slightly better for CPOD compared to MCA, as more clearly seen with the mean-flow results of Fig. 9. Such results ensure the pertinence of CPOD in obtaining coherent structures related to the Reynolds shear stress. Nonetheless, an advantage of MCA is its monotonic convergence, seen in Fig. 9, which is related to the properties of the underlying singular value decomposition.

#### IV. CONCLUSIONS

This work presents a method for the modal decomposition of flows designed to study the cross-covariance between fluctuations, referred to as cross proper orthogonal decomposition, or CPOD. The method allows one to extract from flow databases modes that are optimal in representing the aforementioned cross-covariances, such as Reynolds stresses in turbulent flows. Such modes appear simultaneously in the fields considered in the cross-covariances, a property that distinguishes the present approach from other methods in the literature. This property allows interpreting leading CPOD modes as coherent structures with maximal contribution to the trace of the cross-covariance at hand.

Application of CPOD is exemplified in turbulent channel flow at friction Reynolds number 179, where it ranks modes based on their contribution to the Reynolds shear stress. The use of CPOD modes as a basis allows low-rank reconstructions of the Reynolds stress, and consequently

of the mean velocity profile. This highlights instantaneous flow fluctuations that have significant contribution to the mean flow. As quantities of interest, such as wall friction and heat transfer, are related to mean values, CPOD appears as a valuable method to obtain fluctuating turbulent structures that are directly related to such mean quantities via the Reynolds stresses. Such structures are promising quantities to be targeted by flow control.

The present method is also directly applicable to study generalized Reynolds stresses [14], i.e., nonlinear terms in the equations for the fluctuations (nonzero frequency-wave-number combinations). Such terms involve products of velocity fluctuations taken at different frequencies or wave numbers, which may be dealt by considering separate realization matrices for the wave-number-frequency pair related to the nonlinear interaction of interest. As nonlinear terms have been recently shown to be structured even at higher Reynolds numbers [29], the application of CPOD to generalized Reynolds stresses is a promising direction.

### APPENDIX: SNAPSHOT METHOD FOR MCA

We here derive a snapshot method for MCA, which enables its application with large datasets. With the notation of the present work, we would like to obtain pairs of modes which represent optimally  $\mathbf{P}_{uv} = \mathbf{C}_u \mathbf{X} \mathbf{Y}^H \mathbf{C}_v^H$ . This can be obtained by a singular value decomposition, a well known technique for a low-rank representation of an arbitrary matrix with two orthonormal bases. If one wishes to obtain orthonormal bases satisfying

$$\mathbf{U}^H \mathbf{W}_u \mathbf{U} = \mathbf{I}, \quad (\text{A1a})$$

$$\mathbf{V}^H \mathbf{W}_v \mathbf{V} = \mathbf{I}, \quad (\text{A1b})$$

the appropriate singular value decomposition is

$$\mathbf{W}_u^{1/2} \mathbf{C}_u \mathbf{X} \mathbf{Y}^H \mathbf{C}_v^H (\mathbf{W}_v^{1/2})^H = \hat{\mathbf{U}} \Sigma \hat{\mathbf{V}}^H, \quad (\text{A2})$$

where the Cholesky decomposition of weights  $\mathbf{W}_u = \mathbf{W}_u^{1/2} (\mathbf{W}_u^{1/2})^H$  (and similarly for  $\mathbf{W}_v$ ) is used. Left and right eigenfunctions are obtained as columns of  $\mathbf{U} = \mathbf{W}_u^{-1/2} \hat{\mathbf{U}}$  and  $\mathbf{V} = \mathbf{W}_v^{-1/2} \hat{\mathbf{V}}$ , satisfying the orthonormality conditions of Eq. (A1), and

$$\mathbf{C}_u \mathbf{X} \mathbf{Y}^H \mathbf{C}_v^H = \mathbf{U} \Sigma \mathbf{V}^H, \quad (\text{A3})$$

the desired singular-value decomposition of  $\mathbf{P}_{uv}$ .

For a small number of grid points, Eq. (A2) is directly applicable, but large datasets require recasting the problem with a snapshot method. This can be obtained by writing an eigenvalue problem for  $\mathbf{U}$ ,

$$\mathbf{P}_{uv} \mathbf{W}_v \mathbf{P}_{uv}^H \mathbf{W}_u \mathbf{U} = \mathbf{C}_u \mathbf{X} \mathbf{Y}^H \mathbf{C}_v^H \mathbf{W}_v \mathbf{C}_v \mathbf{Y} \mathbf{X}^H \mathbf{C}_u^H \mathbf{W}_u \mathbf{U} = \mathbf{U} \Sigma^2, \quad (\text{A4})$$

where, considering that  $\mathbf{U}$  is a linear combination of  $\mathbf{C}_u \mathbf{X}$  realizations,

$$\mathbf{U} = \mathbf{C}_u \mathbf{X} \Phi, \quad (\text{A5})$$

leads to an eigenvalue problem for the  $\Phi$  coefficients

$$\mathbf{Y}^H \mathbf{C}_v^H \mathbf{W}_v \mathbf{C}_v \mathbf{Y} \mathbf{X}^H \mathbf{C}_u^H \mathbf{W}_u \mathbf{C}_u \mathbf{X} \Phi = \Phi \Sigma^2. \quad (\text{A6})$$

This eigenvalue problem has dimension given by the number of realizations, and, if this is much lower than the number of degrees of freedom, the present snapshot method is a computationally efficient way to obtain MCA modes. A similar procedure leads to an eigenvalue problem for the right MCA modes, given by

$$\mathbf{X}^H \mathbf{C}_u^H \mathbf{W}_u \mathbf{C}_u \mathbf{X} \mathbf{Y}^H \mathbf{C}_v^H \mathbf{W}_v \mathbf{C}_v \mathbf{Y} \Psi = \Psi \Sigma^2, \quad (\text{A7})$$

where

$$\mathbf{V} = \mathbf{C}_v \mathbf{Y} \Psi. \quad (\text{A8})$$

The  $\Phi$  and  $\Psi$  matrices have columns that are thus right and left eigenvectors, respectively, of

$$\mathbf{A}_{\text{MCA}} = \mathbf{Y}^H \mathbf{C}_v^H \mathbf{W}_v \mathbf{C}_v \mathbf{Y} \mathbf{X}^H \mathbf{C}_u^H \mathbf{W}_u \mathbf{C}_u \mathbf{X}, \quad (\text{A9})$$

and thus satisfy a biorthogonality relation, such that  $\Psi^H \Phi$  and  $\Phi^H \Psi$  are diagonal matrices if all eigenvalues of  $\mathbf{A}_{\text{MCA}}$  are distinct. A convenient choice for normalization is  $\Psi^H \Phi = \Sigma^{-1} = \Phi^H \Psi$ , as will be shown shortly.

The coefficients of an expansion onto MCA modes,  $\Gamma_u$ , may be found following a procedure similar to Sec. II D, via

$$\Gamma_u \Phi = \mathbf{U}^H \mathbf{W}_u \mathbf{C}_u \mathbf{X} \Phi = \mathbf{I}, \quad (\text{A10})$$

and thus

$$\Gamma_u = \Phi^{-1} = \Sigma \Psi^H. \quad (\text{A11})$$

A similar reasoning leads to  $\Gamma_v = \Sigma \Phi^H$ . With such relations, we have  $\mathbf{C}_u \mathbf{X} = \mathbf{U} \Gamma_u$  and  $\mathbf{C}_v \mathbf{Y} = \mathbf{V} \Gamma_v$ , and hence

$$\mathbf{C}_u \mathbf{X} \mathbf{Y}^H \mathbf{C}_v^H = \mathbf{U} \Gamma_u \Gamma_v^H \mathbf{V}^H = \mathbf{U} \Sigma \Psi^H \Phi \Sigma \mathbf{V}^H = \mathbf{U} \Sigma \mathbf{V}^H, \quad (\text{A12})$$

as desired. Notice that  $\Gamma_u \neq \Gamma_v$ , and thus left and right MCA modes have different time evolutions, i.e., they evolve asynchronously, unlike CPOD modes.

As a final remark, although MCA modes so defined only involve the observations extracted using  $\mathbf{C}_u$  and  $\mathbf{C}_v$ , with the present snapshot method it is nonetheless possible to reconstruct modes involving all degrees of freedom of  $\mathbf{X}$  and  $\mathbf{Y}$ , as in CPOD. These are obtained as

$$\mathbf{U}_f = \mathbf{X} \Phi, \quad (\text{A13a})$$

$$\mathbf{V}_f = \mathbf{Y} \Psi. \quad (\text{A13b})$$

In the turbulent channel flow example of Sec. III, whereas  $\mathbf{U}$  and  $\mathbf{V}$  modes would only comprise streamwise and wall-normal velocity components  $u$  and  $v$ , respectively, the full MCA modes  $\mathbf{U}_f$  and  $\mathbf{V}_f$  involve the three velocity components. As such modes are obtained from linear combinations of snapshots, they inherit properties such as zero divergence and nonslip boundary conditions, similar to POD and CPOD modes.

- 
- [1] P. Jordan and T. Colonius, Wave packets and turbulent jet noise, *Annu. Rev. Fluid Mech.* **45**, 173 (2013).
  - [2] J. Jiménez, Coherent structures in wall-bounded turbulence, *J. Fluid Mech.* **842**, P1 (2018).
  - [3] K. Taira, S. L. Brunton, S. T. Dawson, C. W. Rowley, T. Colonius, B. J. McKeon, O. T. Schmidt, S. Gordeyev, V. Theofilis, and L. S. Ukeiley, Modal analysis of fluid flows: An overview, *AIAA J.* **55**, 4013 (2017).
  - [4] G. Berkooz, P. Holmes, and J. Lumley, The proper orthogonal decomposition in the analysis of turbulent flows, *Annu. Rev. Fluid Mech.* **25**, 539 (1993).
  - [5] I. Jolliffe, *Principal Component Analysis*, Springer Series in Statistics (Springer-Verlag, New York, 2002).
  - [6] N. Aubry, P. Holmes, J. Lumley, and E. Stone, The dynamics of coherent structures in the wall region of a turbulent boundary layer, *J. Fluid Mech.* **192**, 115 (1988).
  - [7] J. Baltzer, R. Adrian, and X. Wu, Structural organization of large and very large scales in turbulent pipe flow simulation, *J. Fluid Mech.* **720**, 236 (2013).
  - [8] L. H. O. Hellström, I. Marusic, and A. J. Smits, Self-similarity of the large-scale motions in turbulent pipe flow, *J. Fluid Mech.* **792**, R1 (2016).
  - [9] 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).
  - [10] T. Smith, J. Moehlis, and P. Holmes, Low-dimensional models for turbulent plane Couette flow in a minimal flow unit, *J. Fluid Mech.* **538**, 71 (2005).

- [11] C. Picard and J. Delville, Pressure velocity coupling in a subsonic round jet, *Int. J. Heat Fluid Flow* **21**, 359 (2000).
- [12] P. J. Schmid, Nonmodal stability theory, *Annu. Rev. Fluid Mech.* **39**, 129 (2007).
- [13] B. McKeon, The engine behind (wall) turbulence: Perspectives on scale interactions, *J. Fluid Mech.* **817**, P1 (2017).
- [14] B. J. McKeon and A. S. Sharma, A critical-layer framework for turbulent pipe flow, *J. Fluid Mech.* **658**, 336 (2010).
- [15] S. Beneddine, D. Sipp, A. Arnault, J. Dandois, and L. Lesshafft, Conditions for validity of mean flow stability analysis, *J. Fluid Mech.* **798**, 485 (2016).
- [16] A. V. G. Cavalieri, P. Jordan, and L. Lesshafft, Wave-packet models for jet dynamics and sound radiation, *Appl. Mech. Rev.* **71**, 020802 (2019).
- [17] A. Towne, O. T. Schmidt, and T. Colonius, Spectral proper orthogonal decomposition and its relationship to dynamic mode decomposition and resolvent analysis, *J. Fluid Mech.* **847**, 821 (2018).
- [18] S. Bagheri, D. Henningson, J. Hoepffner, and P. Schmid, Input-output analysis and control design applied to a linear model of spatially developing flows, *Appl. Mech. Rev.* **62**, 020803 (2009).
- [19] B. F. Farrell and P. J. Ioannou, Stochastic forcing of the linearized Navier-Stokes equations, *Phys. Fluids A: Fluid Dyn.* **5**, 2600 (1993).
- [20] A. Towne, A. Lozano-Durán, and X. Yang, Resolvent-based estimation of space-time flow statistics, *J. Fluid Mech.* **883**, A17 (2020).
- [21] A. Zare, M. R. Jovanović, and T. T. Georgiou, Colour of turbulence, *J. Fluid Mech.* **812**, 636 (2017).
- [22] O. T. Schmidt, A. Towne, G. Rigas, T. Colonius, and G. A. Brès, Spectral analysis of jet turbulence, *J. Fluid Mech.* **855**, 953 (2018).
- [23] L. Lesshafft, O. Semeraro, V. Jaunet, A. V. G. Cavalieri, and P. Jordan, Resolvent-based modeling of coherent wave packets in a turbulent jet, *Phys. Rev. Fluids* **4**, 063901 (2019).
- [24] L. I. Abreu, A. V. G. Cavalieri, P. Schlatter, R. Vinuesa, and D. S. Henningson, Resolvent modelling of near-wall coherent structures in turbulent channel flow, *Int. J. Heat Fluid Flow* **85**, 108662 (2020).
- [25] L. I. Abreu, A. V. G. Cavalieri, P. Schlatter, R. Vinuesa, and D. S. Henningson, Spectral proper orthogonal decomposition and resolvent analysis of near-wall coherent structures in turbulent pipe flows, *J. Fluid Mech.* **900**, A11 (2020).
- [26] J. M. Wallace, Quadrant analysis in turbulence research: History and evolution, *Annu. Rev. Fluid Mech.* **48**, 131 (2016).
- [27] J. C. del Alamo and J. Jiménez, Spectra of the very large anisotropic scales in turbulent channels, *Phys. Fluids* **15**, L41 (2003).
- [28] D. Krug, D. Lohse, and R. J. Stevens, Coherence of temperature and velocity superstructures in turbulent Rayleigh-Bénard flow, *J. Fluid Mech.* **887**, A2 (2020).
- [29] P. Morra, P. A. S. Nogueira, A. V. G. Cavalieri, and D. S. Henningson, The colour of forcing statistics in resolvent analyses of turbulent channel flows, *J. Fluid Mech.* **907**, A24 (2021).
- [30] P. A. S. Nogueira, P. Morra, E. Martini, A. V. G. Cavalieri, and D. S. Henningson, Forcing statistics in resolvent analysis: Application in minimal turbulent Couette flow, *J. Fluid Mech.* **908**, A32 (2021).
- [31] O. T. Schmidt, Bispectral mode decomposition of nonlinear flows, *Nonlin. Dyn.* **102**, 2479 (2020).
- [32] C. S. Bretherton, C. Smith, and J. M. Wallace, An intercomparison of methods for finding coupled patterns in climate data, *J. Clim.* **5**, 541 (1992).
- [33] J. Jiménez, Near-wall turbulence, *Phys. Fluids* **25**, 101302 (2013).
- [34] P. Hall and S. Sherwin, Streamwise vortices in shear flows: Harbingers of transition and the skeleton of coherent structures, *J. Fluid Mech.* **661**, 178 (2010).
- [35] J. M. Hamilton, J. Kim, and F. Waleffe, Regeneration mechanisms of near-wall turbulence structures, *J. Fluid Mech.* **287**, 317 (1995).
- [36] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis* (Cambridge University Press, Cambridge, UK, 1994).
- [37] S. Axler, *Linear Algebra Done Right* (Springer, New York, 2015).
- [38] D. Rodríguez, A. V. Cavalieri, T. Colonius, and P. Jordan, A study of linear wavepacket models for subsonic turbulent jets using local eigenmode decomposition of PIV data, *Eur. J. Mech.-B* **49**, 308 (2015).

- [39] V. Theofilis, Advances in global linear instability analysis of nonparallel and three-dimensional flows, *Prog. Aerosp. Sci.* **39**, 249 (2003).
- [40] A. L. van den Wollenberg, Redundancy analysis an alternative for canonical correlation analysis, *Psychometrika* **42**, 207 (1977).
- [41] W. C. Thacker, Principal predictors, *Int. J. Climatol.* **19**, 821 (1999).
- [42] M. Schlegel, B. R. Noack, P. Jordan, A. Dillmann, E. Gröschel, W. Schröder, M. Wei, J. B. Freund, O. Lehmann, and G. Tadmor, On least-order flow representations for aerodynamics and aeroacoustics, *J. Fluid Mech.* **697**, 367 (2012).
- [43] J. F. Gibson, F. Reetz, S. Azimi, A. Ferraro, T. Kreilos, H. Schrobsdorff, M. Farano, A. F. Yesil, S. S. Schütz, M. Culp, and T. M. Schneider, Channelflow 2.0 (see [channelflow.ch](http://channelflow.ch)) (unpublished).
- [44] E. Martini, A. V. Cavalieri, P. Jordan, A. Towne, and L. Lesshafft, Resolvent-based optimal estimation of transitional and turbulent flows, *J. Fluid Mech.* **900**, A2 (2020).
- [45] L. N. Trefethen, *Spectral Methods in MATLAB* (Society for Industrial Mathematics, Philadelphia, 2000), Vol. 10.
- [46] N. Hutchins and I. Marusic, Large-scale influences in near-wall turbulence, *Philos. Trans. R. Soc., A* **365**, 647 (2007).