IOPscience iopscience.iop.org

Home Search Collections Journals About Contact us My IOPscience

Granger causality in wall-bounded turbulence

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2014 J. Phys.: Conf. Ser. 506 012006

(http://iopscience.iop.org/1742-6596/506/1/012006)

View the table of contents for this issue, or go to the journal homepage for more

Download details:

IP Address: 161.24.8.137
This content was downloaded on 28/10/2014 at 19:52

Please note that terms and conditions apply.

Journal of Physics: Conference Series **506** (2014) 012006 doi:10.1088/1742-6596/506/1/012006

Granger causality in wall-bounded turbulence

Gilles Tissot[1], Adrian Lozano–Durán[2], Laurent Cordier[1], Javier Jiménez[2], Bernd R Noack[1]

[1]Institut PPRIME, CEAT 43 route de l'aérodrome, 86036 Poitiers, France
[2]School of Aeronautics, Universidad Politécnica, 28040 Madrid, Spain

E-mail: Gilles.Tissot@univ-poitiers.fr, Adrian@torroja.dmt.upm.es, Laurent.Cordier@univ-poitiers.fr, Jimenez@torroja.dmt.upm.es, Bernd.Noack@univ-poitiers.fr

Abstract. Granger causality is based on the idea that if a variable helps to predict another one, then they are probably involved in a causality relationship. This technique is based on the identification of a predictive model for causality detection. The aim of this paper is to use Granger causality to study the dynamics and the energy redistribution between scales and components in wall-bounded turbulent flows. In order to apply it on flows, Granger causality is generalized for snapshot-based observations of large size using linear-model identification methods coming from model reduction. Optimized DMD, a variant of the Dynamic Mode Decomposition, is considered for building a linear model based on snapshots. This method is used to link physical events and extract physical mechanisms associated to the bursting process in the logarithmic layer of a turbulent channel flow.

1. Introduction

Deduction and induction are the two opposite faces of scientific construction. In a deductive approach, accepted principles are used (for instance physical laws) to reach some conclusion by pure logical reasoning. Observations are then only employed for invalidating hypothesis or theory. At the opposite, with inductive reasoning, we try from specific observations to derive broader generalizations and theories. Causality, at the intersection of philosophy and sciences [1], enters directly in this latter category. In physics, different statistical tools have been developed for analysing causes and effects between phenomena [2]. A common approach for inferring information flow from time series is to compute the cross-correlation between two variables for a range of time lags, and determine whether there exists a peak in the correlation at some non-zero lag. However, this approach can be misleading to infer causation since correlation only indicates if two variables are statistically linked. One possible link may be a direct cause-effect relationship between the two variables, but that cause and effect can be also indirect, or due to a confounding variable i.e. an unseen variable that is also correlated with the two variables. Moreover, since correlation is a symmetric measure, it does not give information on the direction of causality. In classical physics, we can exploit the fact that "causes precede the effects" (principle of time-asymmetry of causation). A dynamical model is another way to include a direction in the variation of the time variable. Indeed, the identification of a predictive model is based on the assumption that the system is dictated by a law, and that the system in the same state will evolve in the same way. In other words, the same causes will produce the same effects (physical determinism). In general, this dynamical law is identified by considering

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

statistics of the observations. Based on a regression of observations (often done through the computation of correlations), a model is determined which links at best variables and their temporal evolutions.

In this paper, Granger causality [3, 4] is considered for detecting causation. This method is one of the earliest statistical tests developed to quantify the temporal-causal effect among time series. It has been successfully employed in the past in economy [5], in neuroscience [6] and in climate modelling [7]. There are mainly two manners of using Granger causality for a given set of data: i) to infer the existence and direction of influence from information in the data; ii) to determine which one of several different hypotheses concerning a physical phenomenon of interest is more likely, given the data and assumptions. The first approach is more exploratory, since it does not rely on a priori specification of dominating events and assumptions about the existence and direction of influence between these events. The crucial point is then the selection of the data to be analysed. Once the data are selected, one needs only apply the Granger causality principle to obtain the causations. This approach is typically used in cognitive neuroscience for determining the interactions of activated brain areas from functional brain imaging [6]. In the second approach, one needs to obtain data relevant to the competing hypotheses before using Granger causality tests to provide clues about the direction of causality. This approach was used in [7] to shed light on a debate about two theories proposed for explaining the relationships between hurricanes and climate. In accordance to the first approach, Granger causality is hereafter used to determine causality relationships between different velocity components and scales in a wall-bounded turbulent flow. Formally, x_2 "Granger-causes" x_1 if information in the past of x_2 helps to predict the future of x_1 with better accuracy than it is possible when considering only information in the past of x_1 itself. Mathematically, Granger causality is based on the identification of linear autoregressive (AR) models for the two variables x_1 and x_2. A rigorous definition of Granger causality is postponed to section 2.

In a wall-bounded turbulent flow, elongated structures of low streamwise velocity, denoted as *streaks*, are present in the so-called buffer layer near the wall [8, 9]. These streaks are known to play an important role in the wall-turbulence phenomena (see [10] for an overview). Indeed, it has been observed in [11] with dye visualisations that the streaks tend to lift-up from the wall, oscillate, and then burst. Optimal perturbations [12, 13] showed that streamwise vortices are amplified in wall-bounded flows with a spacing that is consistent with observed streaks. These vortices are responsible for changing the mean flow and for the creation of the streaks. Reference [14] studied in detail the growth of the spanwise instabilities in sufficiently strong streaks, leading to their breakdown. This process is associated with strong ejections of fluid from the wall, called *bursts*, and to interactions with parts of the flow farther from the wall. These phenomena contribute in large part to changes of the Reynolds stresses. Farther from the wall, the logarithmic layer is also organized into streamwise streaks, although they differ from those in the buffer layer by their larger dimensions and their highly multiscale characteristics [15]. These streaks equally burst [16], and the associated strong intermittent wall-normal fluctuations were interpreted by [17] as a linear amplification that can be modelled by the Orr mechanism. The use of the linear Orr-Sommerfeld equation was effectively able to reproduce some of the burst features observed in fully turbulent simulations [17], especially for the wall-normal velocity. Focusing on that kind of events, we propose to adopt the opposite point of view, and to detect Granger causality by the identification of linear models of representative variables in the fully turbulent flow, rather using a linearized model.

Turbulent flows are characterized by a high level of spatio-temporal complexity that leads to dynamical models with a high number of active degrees of freedom. For this type of flows, the traditional approach of Granger causality that relies on identification of autoregressive models through least-squares method is no longer possible. Since the predictive abilities of the linear identified model is essential for the Granger analysis, particular attention should be given to the

modelling part. *Reduced-order models* have demonstrated their ability in dynamical modelling for flow control [18, 19] where the constraints in terms of accuracy and representativeness are equally strong. In order to adapt the strategy used by Granger [3] to a snapshot-based point of view, *Optimized Dynamic Mode Decomposition* (DMD) recently introduced by [20] will be used as a linear system identification method.

In section 2, Granger causality is first presented in its original form. The method is adapted in section 3 to the case, typical in fluid mechanics, where the size of the snapshots is very large. For that, optimized DMD is introduced in section 3.1 as a way of identifying a linear system based on snapshots. The use of optimized DMD for flow prediction is demonstrated in section 3.2, and the Granger causality formalism is adapted in section 3.3 to the use of an optimized DMD model. Finally, Granger causalities between velocity components of different scales representative of the streaks are determined in section 4 from numerical data of a channel flow.

2. Granger causality

Let x_1 and x_2 be two discrete-time series of length N_t. The idea of Granger causality [3] formalizes Wiener's intuition that x_2 causes x_1 if knowing x_2 helps to predict the future of x_1. More precisely, we will say that x_2 Granger-causes x_1 if the inclusion of past observations of x_2 reduces the prediction error of x_1 in a linear regression model of x_1 and x_2, as compared to a model which includes only previous observations of x_1. Clearly, the notion of Granger causality does not imply true causality. It only implies forecasting ability.

Assume that x_1 can be described by an univariate autoregressive model of order $p < N_t$, or VAR(p) model, given by

$$x_1(t_k) = \sum_{i=1}^{p} a_i x_1(t_{k-i}) + \eta_1(t_k), \tag{1}$$

and also by a bivariate linear regression model defined as

$$x_1(t_k) = \sum_{i=1}^{p} \widetilde{a}_i x_1(t_{k-i}) + \sum_{i=1}^{p} \widetilde{b}_i x_2(t_{k-i}) + \widetilde{\eta}_1(t_k). \tag{2}$$

In (1) and (2), $\{a_i, \widetilde{a}_i, \widetilde{b}_i\}_{i=1}^{p}$ are called the coefficients of the model, and $\eta_1, \widetilde{\eta}_1$ are white noise processes corresponding to the prediction errors for each time series.

For measuring Granger causality, reference [21] introduced the log ratio of the prediction error variances for the two autoregressive models (1) and (2), i.e.

$$\mathcal{F}_{2\to 1} = \ln\left(\frac{\mathrm{var}(\eta_1)}{\mathrm{var}(\widetilde{\eta}_1)}\right). \tag{3}$$

This definition has several motivations. First, since adding information should statistically improve the prediction, the variance of η_1 is greater than or equal to the variance of $\widetilde{\eta}_1$, and $\mathcal{F}_{2\to 1} \geq 0$, as any measure must be. Second, $\mathcal{F}_{2\to 1}$ is a monotonic function with respect to prediction quality, which can be used to compare the magnitude of Granger causality.

Importantly, the concept of Granger causality is easy to extend to vectorial time series [2]. In that case, the set of N_x time series $\boldsymbol{x}(t_k) = (x_1(t_k), \cdots, x_{N_x}(t_k))^T$ is assumed to be adequately described by the multivariate autoregressive (MVAR) model

$$\boldsymbol{x}(t_k) = \sum_{i=1}^{p} A(i)\boldsymbol{x}(t_{k-i}) + \boldsymbol{\eta}(t_k) \tag{4}$$

where $A(i)$, $i = 1, \cdots, p$ are $N_x \times N_x$ matrices containing the elements $A_{jl}(i)$ that describe the linear interaction at time-lag i from $x_j(t_k)$ to $x_l(t_{k-i})$, and $\boldsymbol{\eta}(t_k) = (\eta_1(t_k), \cdots, \eta_{N_x}(t_k))^T$ is a vector of white noise processes. In terms of MVAR coefficients, x_j Granger-causes x_l if at least one off-diagonal element $A_{jl}(i)$ is significantly different from zero. Indeed, a non-zero value of $A_{jl}(i)$ means that the use of the information of $x_l(t_{k-i})$ statistically improves the estimation of $x_j(t_k)$ leading to a decrease of the corresponding variance.

3. Granger causality for high-dimensional time series

The generalization of Granger causality to MVAR model presented in section 2 seems perfectly well suited to the analysis of turbulent flows. Indeed, considering only one-dimensional time signals for causality detection of turbulent events may not be very conclusive, since we do not exploit the spatio-temporal knowledge coming from highly-resolved numerical simulations [16]. However, high dimensionality poses numerous challenges to statistical methods [22]. Indeed, the classical (ordinary least-squares) or more advanced (e.g., [23]) numerical methods used so far to estimate the coefficients of MVAR models are not appropriate for high-dimensional data. For this reason, we propose in this manuscript an alternative way of estimating the linear model necessary for the Granger analysis. This approach is based on a model reduction algorithm introduced recently in the domain of dynamical systems, and called optimized Dynamical Mode Decomposition (DMD) [20]. This method can be viewed as computing, from flow snapshots, eigenvalues and eigenvectors of a linear model that approximates the underlying dynamics, even if those dynamics are nonlinear. As it is the case for all snapshot-based methods, the modelled dynamics is only the part expressed in the snapshots of the original dynamics. As a consequence, the choice of the snapshots is crucial to obtain a physically reliable model.

The standard optimized DMD algorithm will be first presented in section 3.1. The ability of optimized DMD to perform one-step prediction of snapshots will be detailed in section 3.2 and, finally, this ability of forecasting the dynamics will be exploited in section 3.3 to derive criteria for Granger causality.

3.1. Optimized DMD

Suppose that $\boldsymbol{x}(t_k)$ $(k = 1, \ldots, N_t)$ is a set of N_x-dimensional snapshots. Given $p < N_t$, optimized DMD corresponds to seek complex scalars $\{\lambda_j\}_{j=1}^p$ and vectors $\{\boldsymbol{\Phi}_j\}_{j=1}^p$ such that

$$\Gamma = \sum_{k=1}^{N_t} \|\boldsymbol{e}(t_k)\|_2^2 \tag{5}$$

is minimized, where

$$\boldsymbol{x}(t_k) = \sum_{i=1}^{p} \lambda_i^{k-1} \boldsymbol{\Phi}_i + \boldsymbol{e}(t_k). \tag{6}$$

Formally, this equation is equivalent (see Appendix A.1) to the existence of a linear operator $\mathcal{A}$ that connects the field $\boldsymbol{x}(t_k)$ to the subsequent field $\boldsymbol{x}(t_{k+1})$, that is,

$$\boldsymbol{x}(t_{k+1}) = \mathcal{A}\boldsymbol{x}(t_k) + \boldsymbol{e}(t_{k+1}). \tag{7}$$

The expression of $\mathcal{A}$ in terms of $\{\lambda_j\}_{j=1}^p$ and $\{\boldsymbol{\Phi}_j\}_{j=1}^p$ can be deduced directly from (6). Moreover, it can also be demonstrated (Appendix A.1) that $(\boldsymbol{\Phi}_j, \lambda_j)$ are the eigen-elements of $\mathcal{A}$.

To proceed with the description of the optimized DMD algorithm, we define

$$X = (\boldsymbol{x}(t_1), \cdots, \boldsymbol{x}(t_{N_t})) \in \mathbb{C}^{N_x \times N_t} \tag{8}$$

$$\Phi = (\boldsymbol{\Phi}_1, \ldots, \boldsymbol{\Phi}_p) \in \mathbb{C}^{N_x \times p} \tag{9}$$

$$E = (\boldsymbol{e}(t_1), \cdots, \boldsymbol{e}(t_{N_t}))\,, \tag{10}$$

and V the Vandermonde matrix based on λ_j

$$V = \begin{pmatrix} 1 & \lambda_1 & \lambda_1^2 & \cdots & \lambda_1^{N_t-1} \\ 1 & \lambda_2 & \lambda_2^2 & \cdots & \lambda_2^{N_t-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_p & \lambda_p^2 & \cdots & \lambda_p^{N_t-1} \end{pmatrix} \in \mathbb{C}^{p \times N_t}. \tag{11}$$

The optimization problem given by (5) and (6) is then equivalent to find Φ and V such that the Frobenius squared norm $\Gamma = \|E\|_F^2$ is minimized in the model

$$X = \Phi V + E, \tag{12}$$

Of all the choices of Φ, the one which minimizes Γ is

$$\Phi = XV^+ \tag{13}$$

where $V^+ = V^*(VV^*)^{-1}$ is the Moore–Penrose pseudoinverse of V, with V^* the conjugate transpose of V. The residual Γ is therefore given by (see details in Appendix A.3):

$$\Gamma = \|E\|_F^2 = \|X\|_F^2 - \|XV^+V\|_F^2. \tag{14}$$

In [20], a particular global optimization technique was used for computing optimized DMD. The main advantage of global optimization is to seek for a global minimum of the cost function (here Γ). Unfortunately, this class of methods is also very time consuming because of the large number of cost function evaluations that are needed to converge [24]. In this paper, we improved the original algorithm of [20] by using a gradient descent method. For that, the gradient of the cost function Γ with respect to the variations of the eigenvalues λ_j was computed analytically (Appendix A.3). It is formally given by

$$\frac{\partial \Gamma}{\partial \lambda_j} = -\mathrm{tr}\left(X^* X \frac{\partial (V^+V)}{\partial \lambda_j} \right) \tag{15}$$

where tr denotes the trace of a matrix. In the optimized DMD algorithm, we then first determine the eigenvalues $\{\lambda_j\}_{j=1}^p$ with a gradient method, and then employ (13) to determine Φ. The computational cost of the iterative part is then independent of the dimension N_x, which makes it well adapted to high-dimensional snapshots data. Since we know an analytical expression of the gradient, efficient quasi-Newton minimization algorithms can be used leading to an increase of the number of optimization parameters in the minimization problem at constant computational resources. In practice, the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm is employed [25].

3.2. Prediction abilities

By definition of the optimized DMD, the error Γ (see eq. 5) is invariant to permutation of all the snapshots. As a consequence, the residual $\boldsymbol{e}(t)$ determined by optimized DMD cannot be directly used for an analysis of Granger causality for which forecasting ability of the model is necessary. The objective of this section is then to determine a new linear expression for $\boldsymbol{x}$ with an error that can be exploited for the Granger causality. One way of doing that is to first determine an expression of $\boldsymbol{x}$ based on its projection onto the basis Φ as determined by optimized DMD (see

section 3.1), and then to propagate the model for $\boldsymbol{x}$ on one time step. We note $\boldsymbol{x}^{\text{proj}}(t_k)$ the orthogonal projection of $\boldsymbol{x}(t_k)$ onto Φ, and $\nu_j(t_k)$ the corresponding projection coefficients (see Fig. 1) i.e.

$$\boldsymbol{x}(t_k) = \boldsymbol{x}^{\text{proj}}(t_k) + \boldsymbol{e}^{\text{proj}}(t_k) = \sum_{i=1}^{p} \nu_i(t_k)\boldsymbol{\Phi}_i + \boldsymbol{e}^{\text{proj}}(t_k), \tag{16}$$

where $\boldsymbol{e}^{\text{proj}}$ is the error introduced in the projection. The projection coefficients $\nu_i(t_k)$ are determined in Appendix A.2. Since $\boldsymbol{x}(t_k)$ is known for $k = 1, \cdots, N_t$, we can also determine $\boldsymbol{x}^{\text{proj}}(t_k)$ for $k = 1, \cdots, N_t$ with (16).

Knowing $\mathcal{A}$, the linear mapping between subsequent snapshots (see Appendix A.1), a one-step prediction of $\boldsymbol{x}(t_k)$ can be written as

$$\mathcal{A}\boldsymbol{x}(t_k) = \sum_{i=1}^{p} \nu_i(t_k)\boldsymbol{\Phi}_i\lambda_i. \tag{17}$$

Finally, since the best prediction which can be done in the projection subspace spanned by Φ is $\boldsymbol{x}^{\text{proj}}(t_{k+1})$, we define the prediction error $\boldsymbol{e}^{\text{pred}}(t_{k+1})$ as

$$\boldsymbol{e}^{\text{pred}}(t_{k+1}) = \boldsymbol{x}^{\text{proj}}(t_{k+1}) - \sum_{i=1}^{p} \nu_i(t_k)\boldsymbol{\Phi}_i\lambda_i. \tag{18}$$

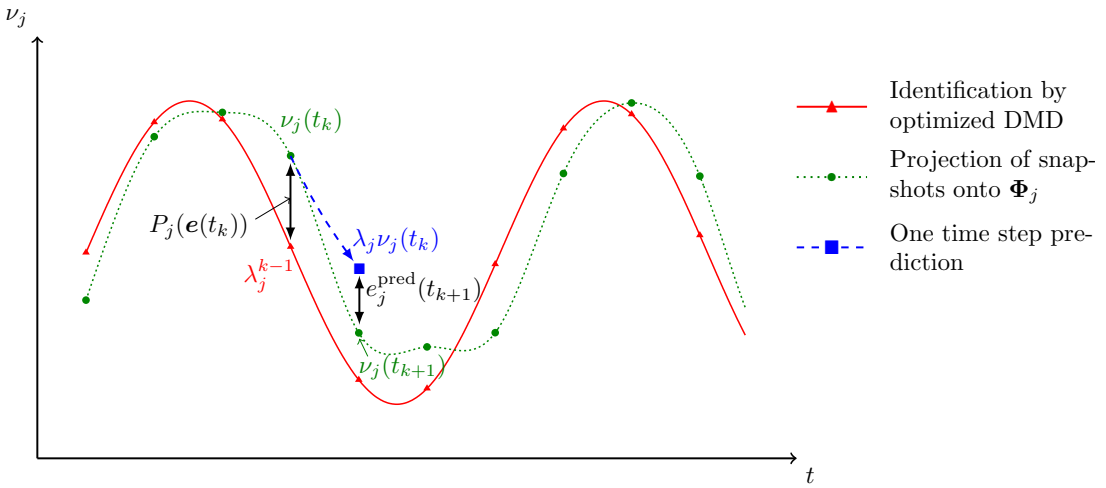


Figure 1. Principle of determination of the prediction error $\boldsymbol{e}^{\text{pred}}$. Linear dynamics identified by optimized DMD (▲), projection of snapshot onto $\boldsymbol{\Phi}_j$ (●) and one time step prediction (■). P_j is the projector onto the vector $\boldsymbol{\Phi}_j$.

The one-step forecast performance is obviously dependent on the model order p. A too low value can lead to a poor representation of the snapshots, whereas a too high value can lead to an "over-fitted" model which can increase the forecast error, as it is demonstrated in [2] for the asymptotic behaviour of estimated MVAR models. The model order may be determined using model selection criteria. The general approach is to choose the value of p which minimizes a criterion that balances the variance accounted for by the model, against the number of coefficients to be estimated. The two most common information criteria are the Akaike (AIC) [26] and the Bayesian Information Criterion (BIC) [27]. These criteria will be used in section 4 for selecting the model order.

3.3. Granger causality using optimized DMD

In sections 3.1 and 3.2, optimized DMD was presented as a method identifying from snapshots a linear model that can be used for one-step prediction. As such, this model can naturally be employed for Granger causality detection.

For the Granger causality, we then consider linear models as given by (18), and assume that λ_i and $\mathbf{\Phi}_i$ have already been determined by optimized DMD. For two time series $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, the Granger causality (see section 2) will then be based on identified linear models given by

$$\begin{cases} \boldsymbol{x}_1^{\text{proj}}(t_{k+1}) = \displaystyle\sum_{i=1}^{p} \nu_{1,i}(t_k)\mathbf{\Phi}_{1,i}\lambda_{1,i} + \boldsymbol{e}_1^{\text{pred}}(t_{k+1}) \\ \boldsymbol{x}_2^{\text{proj}}(t_{k+1}) = \displaystyle\sum_{i=1}^{p} \nu_{2,i}(t_k)\mathbf{\Phi}_{2,i}\lambda_{2,i} + \boldsymbol{e}_2^{\text{pred}}(t_{k+1}), \end{cases} \tag{19}$$

when the prediction is only based on one time series, and

$$\begin{pmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{pmatrix}^{\text{proj}} (t_{k+1}) = \sum_{i=1}^{2p} \widetilde{\nu}_i(t_k) \begin{pmatrix} \widetilde{\mathbf{\Phi}}_{1,i} \\ \widetilde{\mathbf{\Phi}}_{2,i} \end{pmatrix} \widetilde{\lambda}_i + \begin{pmatrix} \widetilde{\boldsymbol{e}}_1^{\text{pred}}(t_{k+1}) \\ \widetilde{\boldsymbol{e}}_2^{\text{pred}}(t_{k+1}) \end{pmatrix} \tag{20}$$

when the prediction is based on both time series. The order of the model (20) is $2p$ to be consistent with the model (2) of the classical formulation of Granger causality (see section 2).

The influence of $\boldsymbol{x}_2$ on $\boldsymbol{x}_1$ ($\mathcal{F}_{2\to1}$), and $\boldsymbol{x}_1$ on $\boldsymbol{x}_2$ ($\mathcal{F}_{1\to2}$) are determined respectively by:

$$\mathcal{F}_{2\to1} = \ln\left(\frac{\displaystyle\int_0^T \|\boldsymbol{e}_1^{\text{pred}}(t)\|^2\,\mathrm{d}t}{\displaystyle\int_0^T \|\widetilde{\boldsymbol{e}}_1^{\text{pred}}(t)\|^2\,\mathrm{d}t}\right) \quad \text{and} \quad \mathcal{F}_{1\to2} = \ln\left(\frac{\displaystyle\int_0^T \|\boldsymbol{e}_2^{\text{pred}}(t)\|^2\,\mathrm{d}t}{\displaystyle\int_0^T \|\widetilde{\boldsymbol{e}}_2^{\text{pred}}(t)\|^2\,\mathrm{d}t}\right). \tag{21}$$

4. Results

In this paper, we propose to analyse the links between different events occurring in a channel flow in the framework of Granger causality. Indeed, although observations seem to link the bursting phenomenon to the evolution of streaks, the mechanisms at the heart of bursts are still not definitely identified, especially in the logarithmic layer. References [11, 17] postulated that the wall-normal velocities are generated by the breakdown of the streamwise-velocity streaks. Here, we suggest to focus on these intermittent events and, for this, we select a set of variables and observations for which we determine their Granger causality following the procedure described in section 3. The objective is to go beyond classical correlations analyses by determining if Granger causality can be useful to understand physical events in the channel flow.

Minimal simulation boxes have demonstrated their ability to generate a single structure in the buffer layer [8] and in the logarithmic layer [15], and also to reproduce statistics of full-size turbulence in the corresponding regions. The data set referred to as W950 in [17] is used in this paper for the Granger analysis. In the following, the streamwise, wall-normal and spanwise directions are denoted by x, y, and z, respectively, and the corresponding velocity components by u, v, and w. The data are obtained from direct numerical simulations of a turbulent channel flow in a minimal box of length $L_x/h = \pi/2$, height $L_y/h = 2$ and width $L_z/h = \pi/4$. These dimensions are considered to be sufficient for isolating typical structures involved in the bursting process in the logarithmic layer [28]. Reference [15] showed that turbulence remains "healthy" roughly below $y \approx 0.3L_z$ corresponding in our computational domain to $y/h \approx 0.25$. The value of the Reynolds number is $Re_\tau = u_\tau h/\nu_{\text{kin}} = 950$, where u_τ is the friction velocity and ν_{kin} is the kinematic viscosity. The incompressible Navier–Stokes equations are integrated in the form

of evolution equations for the wall-normal vorticity, and for the Laplacian of the wall-normal velocity, as in [29], and are computed with a constant mass flux. De-aliased Fourier series are used in the spanwise and streamwise coordinate direction, and a Chebychev polynomial representation is employed in the wall-normal direction. A third-order semi-implicit Runge–Kutta method is used for the time integration, as in [30].

The velocity fields u, v and w are first decomposed as a sum of a temporal mean and their fluctuations, that is, $u(x,y,z,t) = U(y) + u'(x,y,z,t)$ (corresponding equations hold for v and w). We then define $\hat{u}(y,t;k_x,k_z)$ the two-dimensional Fourier transform in the streamwise and spanwise directions of the fluctuating velocity field $u'(x,y,z,t)$ where the wave indices k_x and k_z are related to the wavelengths $\lambda_x = \frac{L_x}{k_x}$ and $\lambda_z = \frac{L_z}{k_z}$, respectively. Since the spanwise direction is homogeneous, the optimized DMD algorithm introduced in section 3.1 is applied to snapshots defined as

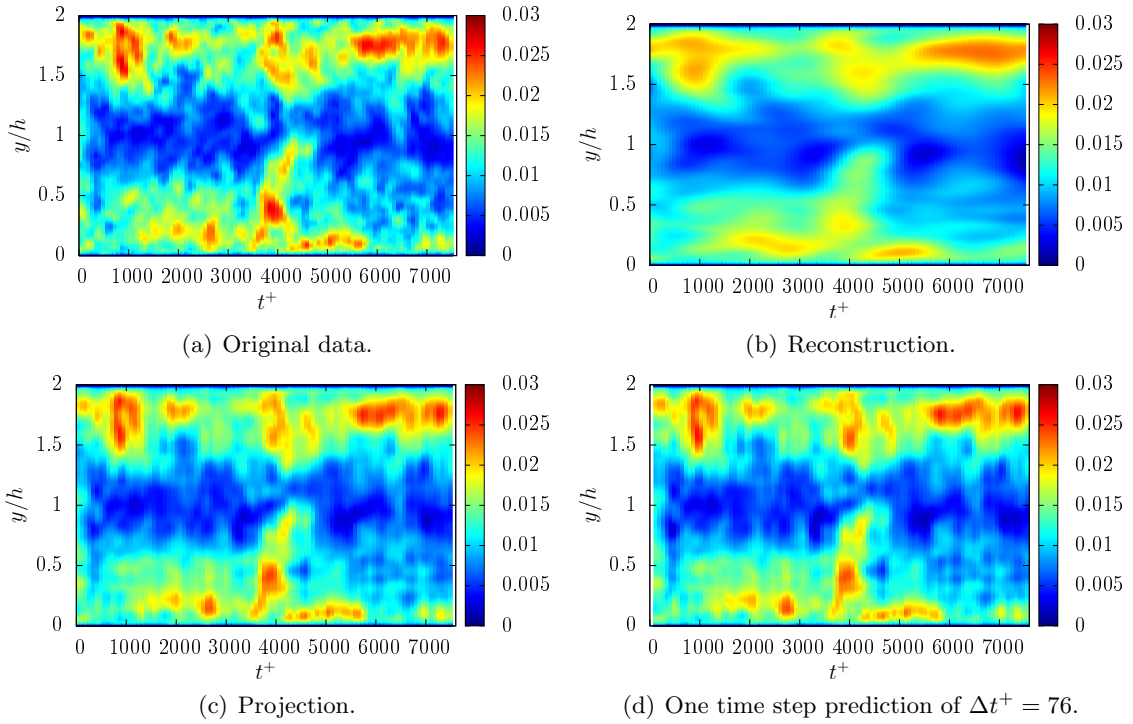
$$\widetilde{u}_{k_x}(y,t_k) = \sqrt{\frac{1}{L_z}\int_{k_z} |\hat{u}(y,t;k_x,k_z)|^2 \, \mathrm{d}k_z} \quad k = 1, \cdots, N_t. \tag{22}$$

These observations correspond to componentwise fluctuating kinetic energy for one streamwise scale k_x allowing us to track the energy cascade along the components and the scales. In order to represent minimal logarithmic-layer structures and their interactions with the outer part of the flow during the bursts, the full wall-normal profile containing $N_y = 385$ points is kept in the analysis. The low-order Fourier modes in the streamwise direction extract the scales associated to the streaks. We then retain for the Granger analysis the snapshots $\widetilde{u}_{k_x}$, $\widetilde{v}_{k_x}$ and $\widetilde{w}_{k_x}$ for $k_x = 0$ and $k_x = 1$, and also the contribution of all the streamwise wave indices from $k_x = 2$ to the last one that we call the "rest". To simplify the notations, these quantities are referred in the following as $\widetilde{u}_0$, $\widetilde{u}_1$, $\widetilde{u}_r$, $\widetilde{v}_0$, $\widetilde{v}_1$, $\widetilde{v}_r$, $\widetilde{w}_0$, $\widetilde{w}_1$ and $\widetilde{w}_r$, respectively.

Optimized DMD models are then used in accordance to section 3 to determine Granger causality between these nine different variables. First, we characterize the ability of an identified linear model determined by optimized DMD to represent, and predict the snapshots. Since it is difficult to imagine that a low-order linear model can predict the behaviour of a turbulent flow for a statistically long time, the identification is restricted to short-term transient events. In this paper, we consider $N_t = 100$ snapshots taken evenly ($\Delta t^+ \approx 76$) over a time window of $T^+ = T\, u_\tau^2/\nu_{\text{kin}} = 7524$ wall time units, typical of a bursting event time scale for the full channel, at $y^+ \approx 1000$ [15]. In optimized DMD, the order p of the linear model is an intrinsic parameter of the method. A compromise between a good level of representativeness of the snapshots and a low-order representation must be done. Indeed, if choosing $p = N_t$ leads to a perfect reconstruction, considering too many modes in the linear expansion may lead numerically to an "over-fitted" model that is not relevant. Following the Bayesian Information Criterion (BIC) [27] which computes a balance between model accuracy and number of parameters, the order of the model is chosen $p = 8$.

The original history of $\widetilde{v}_0$ is represented in Fig. 2(a). For comparison purposes, the approximation of $\widetilde{v}_0$ according to (6) is shown in Fig. 2(b). In that case, only the essential part of the dynamics is described. The projection of the original data onto the optimized DMD basis as given by (16) is illustrated in Fig. 2(c). The approximation of $\widetilde{v}_0$ is rather good, demonstrating that the description based on optimized DMD is qualitatively accurate. Finally, Eq. (18) is used for doing one-time-step prediction. The result, given in Fig. 2(d), should be compared to Fig. 2(c).

At this point, the ability of optimized DMD to represent and predict the observations is characterized. We can now use these linear models to analyse Granger causalities. The magnitude of the interaction between two variables is measured for the nine quantities considered based on criterion (21). $\mathcal{F}_{i\to j}$ for $i, j = 1, \cdots, 9$ determine the level of Granger causality between the variables i and j. These measures are reported in Fig. 3, where i corresponds to



(a) Original data. (b) Reconstruction.

(c) Projection. (d) One time step prediction of $\Delta t^+ = 76$.

Figure 2. Linear system identification of $\widetilde{v}_0$.

the ordinates, and j to the abscissas. By convention, the diagonal elements $\mathcal{F}_{i\to i}$ are equal to 0. We notice immediately that the matrix of Granger causality is not symmetric. The values of the subdiagonal elements are higher than their superdiagonal counterpart, demonstrating globally a direct energy cascade from larger to smaller scales. Moreover, three strong levels of causality are clearly visible, leading us to conclude that: i) $\widetilde{u}_0$ Granger-causes $\widetilde{v}_1$; ii) $\widetilde{u}_0$ Granger-causes $\widetilde{v}_r$; and iii) $\widetilde{u}_1$ Granger-causes $\widetilde{v}_r$. At weaker levels, we also observe the Granger causalities: iv) from $\widetilde{u}_r$ to $\widetilde{v}_r$; v) from $\widetilde{w}_1$ to $\widetilde{v}_1$; and finally, vi) from $\widetilde{w}_r$ to $\widetilde{v}_r$. Remarkably, different scales are involved in the Granger causalities i) to iii), whereas the same scales are participating to the weaker Granger causalities iv) to vi). In terms of wall-turbulence dynamics, the streamwise fluctuating energy for elongated scales, $\widetilde{u}_0$ and $\widetilde{u}_1$, can be interpreted as the presence of meandering streaks, whereas the wall-normal fluctuating energy for the same scales, $\widetilde{v}_0$ and $\widetilde{v}_1$, can be viewed as the sign of ejections and sweeps. Based on these physical interpretations, we can conclude that Granger causalities i) to iii) give strong hints that the dynamical evolution of the streaks causes the ejections and the sweeps. Furthermore, the most likely interpretation of Granger causalities iv) to vi) are the action of oblique vortices.

The variables $\widetilde{u}_0$, $\widetilde{v}_1$, $\widetilde{u}_1$ and $\widetilde{v}_r$ which are involved in the main Granger causalities i) to iii) are represented in Fig. 4. Here, we can see the advantage of a snapshot-based approach that gives an idea of the locations of the involved events. For example, focusing on the $\widetilde{u}_0$–$\widetilde{v}_1$ causality, we can identify at the wall near $t^+ = 4000$ (Fig. 4(a)) an event of strong streamwise fluctuating energy which precedes the wall-normal activity in the outer region of the flow (Fig. 4(b)). Moreover, this event corresponds to streamwise fluctuations that are transformed into wall-normal fluctuations. A plausible interpretation would be that the elongated streamwise fluctuations are transformed in wall-normal fluctuations by the streak disorganization, as observed in [15], creating ejections and sweeps during the bursts. Granger causalities corroborate that the streaks breakdowns cause the bursts, but the way that these streamwise perturbations are transformed into wall-normal fluctuations is not explained by our study.

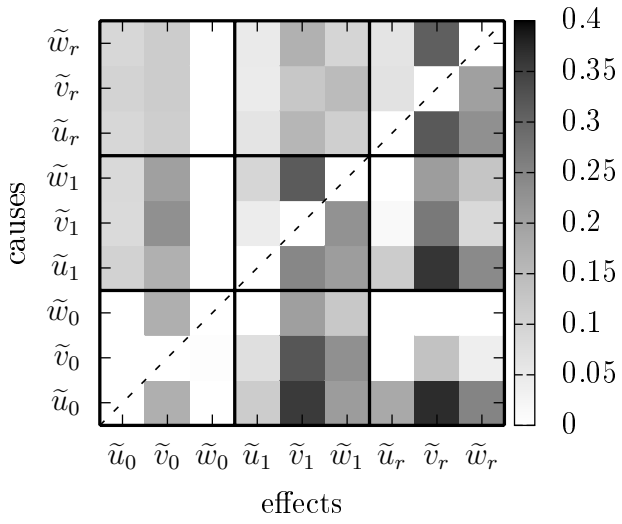
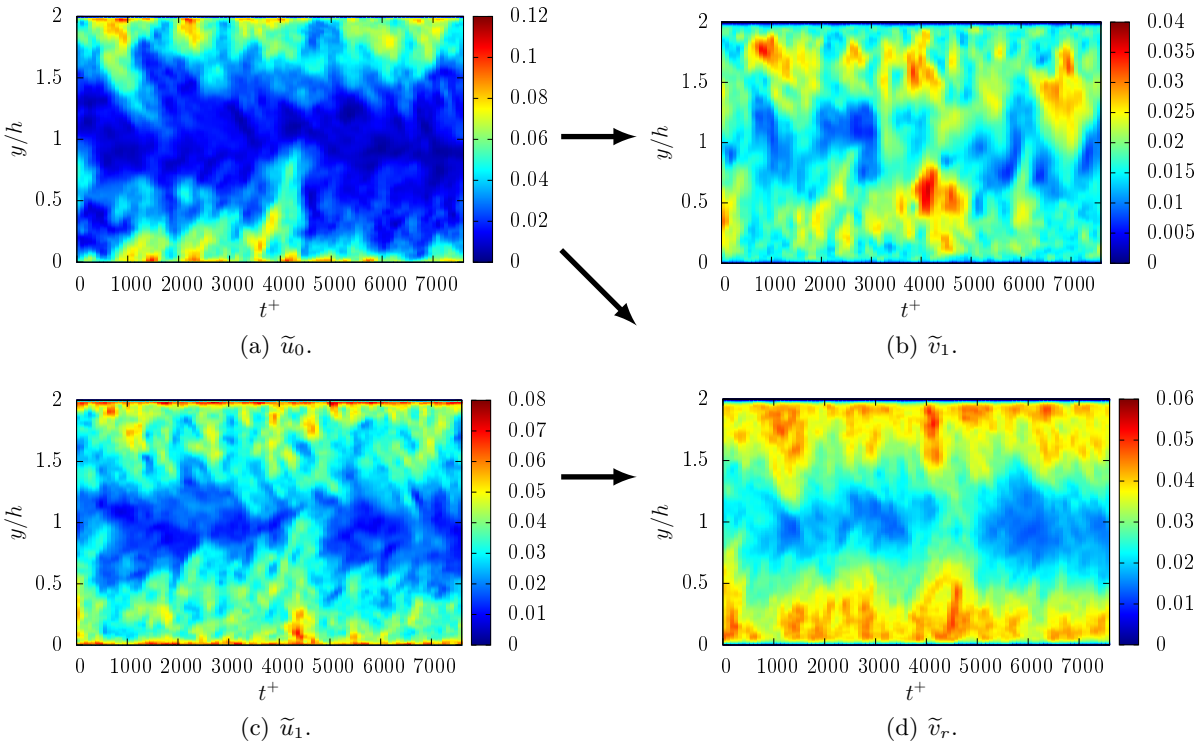


Figure 3. Matrix representation of Granger causalities between $\widetilde{u}_0$, $\widetilde{v}_0$, $\widetilde{w}_0$, $\widetilde{u}_1$, $\widetilde{v}_1$, $\widetilde{w}_1$, $\widetilde{u}_r$, $\widetilde{v}_r$ and $\widetilde{w}_r$ quantified by the prediction improvement $\mathcal{F}_{i\to j}$ as defined in (21). The ordinate i Granger-causes the abscissa j. Heavier lines separate the different scales, whereas the dashed diagonal separates the forward (subdiagonal) and backward (superdiagonal) energy cascade.



(a) $\widetilde{u}_0$. (b) $\widetilde{v}_1$.

(c) $\widetilde{u}_1$. (d) $\widetilde{v}_r$.

Figure 4. Variables involved in the main Granger causalities. The arrows "$\to$" indicate the direction of causality.

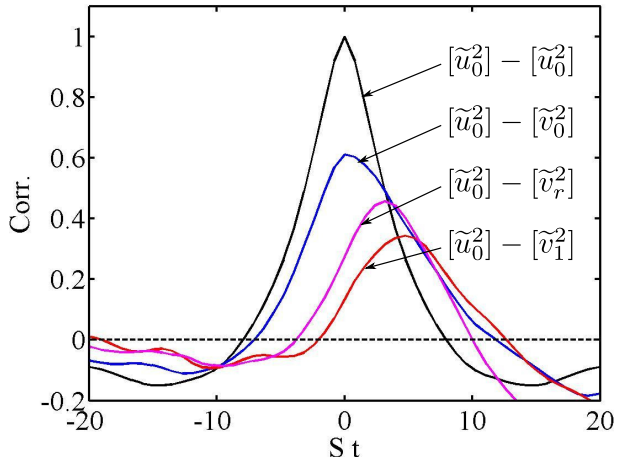


Figure 5. Temporal correlations between $[\widetilde{u}_0^2]$ and $[\widetilde{u}_0^2]$ (black), $[\widetilde{u}_0^2]$ and $[\widetilde{v}_0^2]$ (blue), $[\widetilde{u}_0^2]$ and $[\widetilde{v}_r^2]$ (magenta), and $[\widetilde{u}_0^2]$ and $[\widetilde{v}_1^2]$ (red). Time t is scaled with S, the mean shear averaged over the band from $y_1/h = 0.2$ to $y_2/h = 0.3$.

Those causality behaviours are confirmed by temporal correlations of flow variables integrated over the minimal box corresponding to the dataset W950. For example, $[\widetilde{u}_0^2]$ is computed by averaging the squares of $\widetilde{u}_0$ over a band of y varying from $y/h = 0.2$ to $y/h = 0.3$. Temporal correlations are then calculated for each variable following the method described in [17]. Figure 5 represents the temporal autocorrelation function of $[\widetilde{u}_0^2]$, and the correlations between $[\widetilde{u}_0^2]$ and $[\widetilde{v}_0^2]$, $[\widetilde{v}_1^2]$ and $[\widetilde{v}_r^2]$, respectively. We observe that the level of correlations is strong, and that the correlations of $[\widetilde{u}_0^2]$ with $[\widetilde{v}_r^2]$ and $[\widetilde{v}_1^2]$ are clearly shifted in time suggesting that they are caused by $\widetilde{u}_0$, which is in accordance with the results of Granger causality. We also note that the peak of correlation for $[\widetilde{u}_0^2] - [\widetilde{v}_r^2]$ precedes the corresponding peak for $[\widetilde{u}_0^2] - [\widetilde{v}_1^2]$. In addition, this result is corroborated by a stronger value of Granger causality from $\widetilde{u}_0$ to $\widetilde{v}_r$ than from $\widetilde{u}_0$ to $\widetilde{v}_1$ (Fig. 3), suggesting a more direct influence. All these results indicate that smaller scales are influenced before and more strongly than the larger ones. This observation could be related to Fig. 4, where we notice that $\widetilde{v}_r$, which is associated to smaller scales, is stronger close to the walls than $\widetilde{v}_1$ (see Fig. 4(b)), which is stronger in the outer-region.

In summary, streamwise velocities of low-order streamwise Fourier modes Granger-cause wall-normal velocities of higher-order modes. Alternatively, we can say that "the knowledge of the streaks breakdown helps to predict the bursting process". Moreover, a forward energy redistribution from streaks instabilities to ejections and sweeps was found. This study gives supplementary clues to the fact that strong ejections of fluid are a consequence of the breakdown of the streaks.

5. Conclusion

The capabilities of Granger causality to study physical mechanisms occurring in a turbulent channel flow were illustrated for data obtained from direct numerical simulations. The initial objectives were to improve our understanding of wall-bounded turbulence by studying the energy redistribution between elongated scales in the logarithmic layer. The paper shows that characterizing the causal relationship between significant events is an important but non-trivial aspect of understanding the behaviour of channel flow. As a post-processing approach, the choice of the observations variables is essential in the application of Granger causality. For maximizing our chance of obtaining physically relevant conclusions, we need to choose very carefully the data, especially in turbulence, where the dynamics is complex and the size of the

snapshots is large. Another difficulty of wall-bounded flows is the lack of rigorous definitions of streaks and bursts which prevent us so far to re-discuss in terms of Granger causality the Self-Sustaining Process proposed in [31]. On the other hand, we may conclude that, compared to pure correlation studies, the use of a predictive model confers to Granger causality a stronger basis for determining causality between events. This advantage is balanced by a strong modelling effort. Indeed, in turbulence, reduced-order modelling is still a challenge if a predictive behaviour is expected. In the present study, we show that, when the snapshots are high-dimensional, the linear models at the heart of the Granger causality framework can be determined by optimized DMD. Finally, it should be noted that the predictive causality as defined by Granger, though helpful in understanding feedbacks and interactions in complex systems, does not necessarily imply true causality.

The results of our Granger analysis, consistent with correlations studies, highlight causality links between streaks breakdown and wall-normal activity probably transferred by shear deformation. Besides, the role of active oblique vortices in the wall dynamics has been detected by Granger causality. The successful use of linear systems for the identification does not imply that the detected mechanism is linear. However, this fact is in line with the assumption that the linear Orr mechanism is involved in the bursting process.

In future work, the authors will apply the Granger analysis to better selected data and more complex flows. Moreover, sensitivity tests will be also performed by adding small amount of noise to the data to prove that the values of Granger causalities are sufficiently robust.

Acknowledgments
This work was funded in part by the Multiflow program of the European Research Council.

Appendix

Appendix A.1. Identification of the linear mapping $\mathcal{A}$ between subsequent snapshots
Based on the modes $\mathbf{\Phi}_i$ and corresponding amplitudes λ_i determined by optimized DMD (see section 3.1), the best reconstruction of the snapshots $\{\boldsymbol{x}(t_k)\}_{k=1}^{N_t}$ that can be obtained are given by

$$\check{\boldsymbol{x}}(t_k) = \sum_{i=1}^{p} \mathbf{\Phi}_i \lambda_i^{k-1} = \Phi \boldsymbol{\lambda}_k, \tag{A.1}$$

where $\boldsymbol{\lambda}_k = (\lambda_1^{k-1}, \ldots, \lambda_p^{k-1})^T$. The objective of this section is to determine the expression of the linear mapping $\mathcal{A}$ that connects the field $\check{\boldsymbol{x}}(t_k)$ to the subsequent field $\check{\boldsymbol{x}}(t_{k+1})$, that is,

$$\check{\boldsymbol{x}}(t_{k+1}) = \mathcal{A}\, \check{\boldsymbol{x}}(t_k). \tag{A.2}$$

The one-step prediction of (A.1) can be written as:

$$\check{\boldsymbol{x}}(t_{k+1}) = \sum_{i=1}^{p} \mathbf{\Phi}_i \lambda_i^{k} = \sum_{i=1}^{p} \mathbf{\Phi}_i \lambda_i \lambda_i^{k-1} = \Phi \Lambda \boldsymbol{\lambda}_k \tag{A.3}$$

with $\Lambda = \mathrm{diag}\,(\lambda_1, \ldots, \lambda_p)$. To proceed, we would like to introduce $\check{\boldsymbol{x}}(t_k)$ on the RHS of (A.3). For this, we deduce from (A.1) that $\check{\boldsymbol{x}}$ is spanned by Φ. We then conclude immediately that $P\check{\boldsymbol{x}}(t_k) = \check{\boldsymbol{x}}(t_k)$ where P is the orthogonal projector onto the space spanned by Φ which is introduced in section Appendix A.2, that is,

$$P\check{\boldsymbol{x}}(t_k) = \Phi \left(\Phi^* \Phi\right)^{-1} \Phi^* \check{\boldsymbol{x}}(t_k) = \check{\boldsymbol{x}}(t_k) = \Phi \boldsymbol{\lambda}_k$$

or, by identification, $\boldsymbol{\lambda}_k = \left(\Phi^* \Phi\right)^{-1} \Phi^* \check{\boldsymbol{x}}(t_k)$. Plugging this expression into (A.3), we obtain that

$$\check{\boldsymbol{x}}(t_{k+1}) = \Phi \Lambda \left(\Phi^* \Phi\right)^{-1} \Phi^* \check{\boldsymbol{x}}(t_k)$$

i.e.

$$\mathcal{A} = \Phi\Lambda\left(\Phi^*\Phi\right)^{-1}\Phi^*. \tag{A.4}$$

This result justifies the assertion that optimized DMD can be viewed as a way to identify a linear dynamics representing optimally a set of snapshots. Moreover, we can deduce from (A.4) that

$$\mathcal{A}\Phi = \Phi\Lambda\left(\Phi^*\Phi\right)^{-1}\Phi^*\Phi = \Phi\Lambda, \tag{A.5}$$

demonstrating that $(\boldsymbol{\Phi}_j, \lambda_j)$ are eigen-elements of $\mathcal{A}$.

Finally, consider that the linear mapping $\mathcal{A}$ is applied to any snapshot $\boldsymbol{x}(t_k)$, we have

$$\begin{aligned}\mathcal{A}\boldsymbol{x}(t_k) &= \Phi\Lambda\left(\Phi^*\Phi\right)^{-1}\Phi^*\boldsymbol{x}(t_k) = \Phi\Lambda\boldsymbol{\nu}(t_k)\\ &= \sum_{i=1}^{p}\boldsymbol{\Phi}_i\nu_i(t_k)\lambda_i.\end{aligned} \tag{A.6}$$

The application of $\mathcal{A}$ to $\boldsymbol{x}$ is then equivalent to the multiplication of each component of the optimized DMD expansion by the corresponding optimized DMD eigenvalue.

Appendix A.2. Projection

Let P be the orthogonal projector onto the space spanned by Φ. By definition (see [32]), we have for all $\boldsymbol{x}$

$$P\boldsymbol{x}(t_k) = \sum_{i=1}^{p}\nu_i(t_k)\boldsymbol{\Phi}_i = \Phi\boldsymbol{\nu}(t_k), \tag{A.7}$$

where $P = \Phi(\Phi^*\Phi)^{-1}\Phi^*$. By identification, we then find immediately that

$$\boldsymbol{\nu}(t_k) = (\Phi^*\Phi)^{-1}\Phi^*\boldsymbol{x}(t_k). \tag{A.8}$$

Appendix A.3. Optimized DMD

In this section, we determine the analytical expression of the gradient of $\Gamma = \|E\|_F^2$ with respect to the variation of the DMD eigenvalues λ_j. First, we simplify the expression of Γ using the trace expansion of the Frobenius norm. We find that

$$\begin{aligned}\Gamma &= \|E\|_F^2 = \|X - \Phi V\|_F^2\\ &= \|X - XV^+V\|_F^2 = \|X(\mathbb{I} - V^+V)\|_F^2\\ &= \mathrm{tr}\left(X(\mathbb{I} - V^+V)(\mathbb{I} - V^+V)^*X^*\right)\\ &= \mathrm{tr}\left(X(\mathbb{I} - V^+V - (V^+V)^* + V^+V(V^+V)^*)X^*\right).\end{aligned} \tag{A.9}$$

Since V^+V is a projection and is self-adjoint, we have

$$\begin{aligned}\Gamma &= \mathrm{tr}\left(X(\mathbb{I} - V^+V)X^*\right)\\ &= \mathrm{tr}\left(X(\mathbb{I} - V^+V(V^+V)^*)X^*\right)\\ &= \mathrm{tr}\left(XX^*\right) - \mathrm{tr}\left((XV^+V)(XV^+V)^*\right)\\ &= \|X\|_F^2 - \|XV^+V\|_F^2.\end{aligned} \tag{A.10}$$

The gradient of Γ with respect to λ_j is then given by:

$$\begin{aligned}\frac{\partial \Gamma}{\partial \lambda_j} &= -\frac{\partial}{\partial \lambda_j}\left(\mathrm{tr}\left(X(V^+V)(V^+V)^*X^*\right)\right) \\ &= -\mathrm{tr}\left(X\frac{\partial (V^+V)}{\partial \lambda_j}X^*\right) \\ &= -\mathrm{tr}\left(X^*X\frac{\partial (V^+V)}{\partial \lambda_j}\right).\end{aligned} \tag{A.11}$$

The Gramian matrix X^*X is computed once for all at the beginning of the optimization procedure for reducing the computational cost. Moreover, since $V^+ = V^*(VV^*)^{-1}$, we have to determine the derivative of a product of matrices with respect to a parameter μ, i.e.

$$\frac{\partial (AB)}{\partial \mu} = \frac{\partial A}{\partial \mu}B + A\frac{\partial B}{\partial \mu}, \tag{A.12}$$

and the derivative of the inverse of a matrix A

$$\frac{\partial A^{-1}}{\partial \mu} = -A^{-1}\frac{\partial A}{\partial \mu}A^{-1}. \tag{A.13}$$

Combining all these expressions, we obtain:

$$\frac{\partial (V^+V)}{\partial \lambda_j} = \frac{\partial V^*}{\partial \lambda_j}\left(VV^*\right)^{-1}V - V^*\left(VV^*\right)^{-1}\left(\frac{\partial V}{\partial \lambda_j}V^* + V\frac{\partial V^*}{\partial \lambda_j}\right)\left(VV^*\right)^{-1}V + V^*\left(VV^*\right)^{-1}\frac{\partial V}{\partial \lambda_j} \tag{A.14}$$

where the gradient of V with respect to λ_j is given analytically by

$$\frac{\partial V}{\partial \lambda_j} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & 2\lambda_j & \cdots & (N-1)\lambda_j^{N_t-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}. \tag{A.15}$$

References

[1] McKay Illari P, Russo F and Williamson J (eds) 2011 *Causality in the Sciences* (Oxford University Press)
[2] Lütkepohl H 2007 *New Introduction to Multiple Time Series Analysis* 2nd ed (Berlin: Springer)
[3] Granger C W J 1969 Investigating causal relations by econometric models and cross-spectral methods *Econometrica* 424–38
[4] Seth A K 2010 A MATLAB toolbox for Granger causal connectivity analysis *J. Neurosci. Meth.* **186** 262–73
[5] Granger C W J, Huang B–N and Yang C–W 2000 A bivariate causality between stock prices and exchange rates: evidence from recent Asian flu *Q. Rev. Econ. Financ.* **40** 337–54
[6] Roebroeck A, Formisano E and Goebel R 2005 Mapping directed influence over the brain using Granger causality and fMRI *Neuroimage* **25** 230–42
[7] Elsner J B 2007 Granger causality and atlantic hurricanes *Tellus A* **59** 476–85
[8] Jiménez J and Moin P 1991 The minimal flow unit in near-wall turbulence. *J. Fluid Mech.* **225** 213–40
[9] Waleffe F 2001 Exact coherent structures in channel flow *J. Fluid Mech.* **435** 93–102
[10] Panton R L 2001 Overview of the self-sustaining mechanisms of wall turbulence *Prog. Aerosp. Sci.* **37** 341–83
[11] Kline S J, Reynolds W C, Schraub F A and Runstadler P W 1967 The structure of turbulent boundary layers *J. Fluid Mech.* **30** 741–73
[12] Butler K M and Farrell B F 1993 Optimal perturbations and streak spacing in wall-bounded turbulent shear flow *Phys. Fluids A* **5** 774–77

[13] Del Álamo J C and Jiménez J 2006 Linear energy amplification in turbulent channels *J. Fluid Mech.* **559** 205–13
[14] Schoppa W and Hussain F 2002 Coherent structure generation in near-wall turbulence *J. Fluid Mech.* **453** 57–108
[15] Flores O and Jiménez J 2010 Hierarchy of minimal flow units in the logarithmic layer *Phys. Fluids* **22** 071704
[16] Jiménez J 2013 Near-wall turbulence *Phys. Fluids* **25** 101302
[17] Jiménez J 2013 How linear is wall-bounded turbulence? *Phys. Fluids* **25** 110814
[18] Ravindran S S 2007 Optimal boundary feedback flow stabilization by model reduction *Comput. Method. Appl. M.* **196** 2555–69
[19] Bergmann M, Cordier L and Brancher J P 2005 Optimal rotary control of the cylinder wake using Proper Orthogonal Decomposition reduced-order model *Phys. Fluids* **17** 097101
[20] Chen K K, Tu J H and Rowley C W 2012 Variants of Dynamic Mode Decomposition: Boundary Condition, Koopman, and Fourier Analyses *J. Nonlinear Sci.* **22** 887–915
[21] Geweke J 1982 Measurement of linear dependence and feedback between multiple time series *J. Am. Stat. Assoc.* **77** 304–13
[22] Fan J, Lv J and Qi L 2011 Sparse high-dimensional models in economics *Annu. Rev. Econ.* **3** 291–317
[23] Schlögl A 2006 A comparison of multivariate autoregressive estimators *Signal Process.* **86** 2426–29
[24] Auger A, Hansen N, Zerpa J M P, Ros R and Schoenauer M 2009 Empirical comparisons of several derivative free optimization algorithms *Acte 9-ième Colloq. Nat. Calcul Structures, CSMA* (Giens, France)
[25] Gilbert J C and Lemaréchal C 2009 The module M1QN3 – Version 3.3 *INRIA Rocquencourt & Rhône-Alpes*
[26] Akaike H 1974 A new look at the statistical model identification *IEEE Trans. Autom. Control* **19** 716–23
[27] Schwarz G 1978 Estimating the dimension of a model *Ann. Stat.* **6** 461–64
[28] Jiménez J 2012 Cascades in wall-bounded turbulence *Annu. Rev. Fluid Mech.* **44** 27–45
[29] Kim J, Moin P and Moser R 1987 Turbulence statistics in fully developed channel flow at low Reynolds number *J. Fluid Mech.* **177** 133–66
[30] Spalart P R, Moser R D and Rogers M M 1991 Spectral method for the Navier–Stokes equations with one infinite and two periodic dimensions *J. Comput. Phys.* **96** 297–324
[31] Waleffe F 1997 On a self-sustaining process in shear flows *Phys. Fluids* **9** 883–900
[32] Meyer C 2000 *Matrix Analysis and Applied Linear Algebra: book and solutions manual* (SIAM)