C. R. Mecanique 342 (2014) 410–416

Contents lists available at ScienceDirect

Comptes Rendus Mecanique

www.sciencedirect.com

Separation flow control

Model reduction using Dynamic Mode Decomposition

Réduction de modèle par décomposition en modes dynamiques

Gilles Tissot *, Laurent Cordier, Nicolas Benard, Bernd R. Noack

PPRIME Institute, CEAT, 43, route de l'Aérodrome, 86000 Poitiers, France

ARTICLE INFO

Article history:
Received 6 September 2013
Accepted 2 December 2013
Available online 10 June 2014

Keywords:
Dynamic Mode Decomposition
Reduced-order model
Cylinder wake

Mots-clés :
Décomposition en modes dynamiques
Réduction de modèle
Sillage de cylindre

ABSTRACT

Dynamic Mode Decomposition (DMD) is a recent post-processing technique that extracts from snapshots dynamic relevant information for the flow. Without explicit knowledge of the dynamical operator, the DMD algorithm determines eigenvalues and eigenvectors of an approximate linear model. The ability of DMD to extract dynamically relevant features of the flow predispose it for building a representative reduced-order subspace from the data and for deriving a reduced-order model. The use of the DMD for reduced-order modelling will be investigated in this paper on experimental flow data of a cylinder wake.

RÉSUMÉ

La décomposition en modes dynamiques (DMD) est une technique récente de post-traitement qui extrait des informations liées à la dynamique de l'écoulement à partir d'une séquence de snapshots. Sans connaissance explicite de l'opérateur dynamique, l'algorithme DMD détermine les valeurs et vecteurs propres d'un modèle linéaire approché. La capacité de la DMD à extraire des structures pertinentes en termes de dynamique de l'écoulement la prédispose à la construction, à partir de données, d'un sous-espace de dimension réduite représentatif et au développement d'un modèle de dimension réduite. L'utilisation de la DMD pour la réduction de modèle sera étudiée dans ce papier sur des données expérimentales d'un écoulement de sillage autour d'un cylindre.

1. Introduction

For a turbulent flow, the number of active degrees of freedom is so important that a preliminary step of model reduction is often necessary for having a chance to understand the flow physics or to derive a control strategy. The general objective of model reduction is to extract first, from physical insights or mathematical tools, the building blocks—called modes—that play a dominant role in terms of modelling, and then to derive a dynamical model for the time evolution of the system. As we will see in the following, these two steps can be considered independently or jointly depending on the technique

* Corresponding author.
E-mail addresses: Gilles.Tissot@univ-poitiers.fr (G. Tissot), Laurent.Cordier@univ-poitiers.fr (L. Cordier), Nicolas.Benard@univ-poitiers.fr (N. Benard), Bernd.Noack@univ-poitiers.fr (B.R. Noack).

http://dx.doi.org/10.1016/j.crme.2013.12.011
1631-0721/

used for extracting the modes. From this perspective, the idea of model-based flow control is to exploit the knowledge of a dynamical model for reacting in real-time to the modifications encountered by the flow. Reduced-order models (ROMs) are then well adapted for developing an efficient control strategy. However, finding the appropriate basis for representing the flow in a low-dimensional space is strongly related to a given objective. Indeed, it is somewhat different for a flow to understand the instability mechanisms, to educe the coherent structures mainly responsible for the energy or to represent the non-linear dynamics. Moreover, different strategies can be used to determine the time evolution propagator of the flow states. The dynamical evolution can be intrinsically linked to the method used for determining the modes, or can be classically obtained by Galerkin projection of the governing equations onto the set of modes defining the low-dimensional subspace.

Reduced-order models based on Proper Orthogonal Decomposition (POD) are the most commonly used as the POD modes are optimal in terms of energy content. Despite this property, POD-based ROM is well known to be inaccurate, essentially due to truncation errors, if the model is not improved. Since the energy content is important but is not sufficient in general to catch the dynamical behaviour, we are focusing in this paper on a procedure recently introduced by Schmid [1] called Dynamic Mode Decomposition (DMD). This algorithm was proposed as a method able to extract dynamically relevant flow features from time-resolved experimental or numerical data. It generalises the global stability modes and approximates the eigen-elements of the Koopman operator [2,3]. Our objective is to derive a ROM which will inherit the good dynamical properties of the projection basis.

In Section 2.1, the classical DMD algorithm is presented, followed in Section 2.2 by a discussion of the crucial point of modes' selection. In Section 3, the optimized DMD algorithm is then introduced and an improvement of the original algorithm based on the use of a gradient method is shortly described. Section 4 then presents two ways of predicting the temporal behaviour of the flow outside the time horizon of the snapshots, and ends with the introduction of a data assimilation formalism for combining the two approaches. Finally, in Section 5, we present results obtained on PIV data for a cylinder wake in turbulent regime.

2. Dynamic mode decomposition

2.1. General description

The data is represented in the form of a snapshot sequence, given by a matrix V_1^N defined as

$$V_1^N = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_N) \in \mathbb{R}^{N_x \times N} \tag{1}$$

where $\boldsymbol{v}_i$ is the ith snapshot. In (1), the subscript 1 denotes the first member of the sequence, while the superscript N denotes the last entry in the sequence. Here, we consider the temporal framework of DMD, and assume that the snapshots are separated by a constant sampling time Δt. The DMD algorithm is built on two main assumptions. The first hypothesis is that there exists a linear operator $\mathcal{A}$ to step forward in time the snapshots. Since V_1^N is finite-dimensional, this operator is written as a matrix $A \in \mathbb{R}^{N_x \times N_x}$ such that

$$\boldsymbol{v}_{i+1} = A\boldsymbol{v}_i, \quad \text{for } i = 1, \cdots, N-1 \tag{2}$$

It follows that the subspace spanned by the data set $V_1^N = (\boldsymbol{v}_1, A\boldsymbol{v}_1, \ldots, A^{N-1}\boldsymbol{v}_1)$ corresponds to the Nth Krylov subspace $\mathcal{K}_N(A, \boldsymbol{v}_1)$ generated by A from $\boldsymbol{v}_1$. The goal of DMD is to determine eigenvalues and eigenvectors of A but without first determining A. As such, DMD can be interpreted as an extension of the classical Arnoldi algorithm used to determine eigen-elements of large size problems [4]. In the Arnoldi algorithm, the knowledge of A is exploited to determine an orthonormal basis for the projection subspace of the Rayleigh–Ritz procedure. In the DMD algorithm, the basis of the projection subspace is determined with a "matrix-free" point of view by considering that only snapshots obtained from a time-stepper are available. The matrix A is no longer necessary but the price is an ill-conditioning of the procedure. Since we are interested by eigenvalues of A, we are searching for the spectrum of a similarity matrix of the Galerkin projection of A onto the subspace spanned by the snapshots. When the number of snapshots of the sequence V_1^N increased, it is reasonable to assume that, beyond a given number of snapshots, $\boldsymbol{v}_i$ becomes linearly dependent. The second hypothesis is then to consider that the Nth iterate writes as a linear combination of the previous iterates, i.e.

$$\begin{aligned} \boldsymbol{v}_N &= c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_{N-1}\boldsymbol{v}_{N-1} + \boldsymbol{r} \\ &= V_1^{N-1}\boldsymbol{c} + \boldsymbol{r} \end{aligned} \tag{3}$$

where $\boldsymbol{c}^{\mathrm{T}} = (c_1, c_2, \cdots, c_{N-1})^{\mathrm{T}}$ and $\boldsymbol{r} \in \mathbb{R}^{N_x}$ is the residual vector. A straightforward calculation leads to

$$AV_1^{N-1} = V_2^N = V_1^{N-1}C + \boldsymbol{r}\boldsymbol{e}_{N-1}^{\mathrm{T}} \tag{4}$$

where $\boldsymbol{e}_i$ is the ith Euclidean unitary vector of length $(N-1)$ and C the companion matrix associated with $\boldsymbol{c}$. C is uniquely defined by the coefficients c_i, which may be found using the Moore–Penrose pseudo-inverse of V_1^{N-1}, i.e. $\boldsymbol{c} = (V_1^{N-1})^+\boldsymbol{v}_N = ((V_1^{N-1})^* V_1^{N-1})^{-1}(V_1^{N-1})^*\boldsymbol{v}_N$.

Let $(\boldsymbol{y}_j, \lambda_j)$ be the jth eigen-elements of C, it can be easily proved that $(\boldsymbol{\Phi}_j = V_1^{N-1}\boldsymbol{y}_j, \lambda_j)$ are approximated eigen-elements of A, the so-called Ritz eigenvectors and eigenvalues. Clearly, $\boldsymbol{\Phi}_j$ is defined up to a constant, so we introduce $\widetilde{\boldsymbol{\Phi}}_j$ physically scaled version of $\boldsymbol{\Phi}_j$. The value of the residual $\boldsymbol{r}$ is a good measure of the approximation i.e. of the success of the DMD algorithm. We will see in Section 2.2 that these Ritz eigenvalues can be used to determine the frequency and the growth rate of the linear process. Properties of the eigen-elements of the companion matrix show that the Ritz eigenvectors span the original data. We then have the relation

$$\boldsymbol{v}_k = \sum_{j=1}^{N-1} \lambda_j^{k-1} \widetilde{\boldsymbol{\Phi}}_j \quad k = 1, \cdots, N-1 \tag{5}$$

where $\widetilde{\boldsymbol{\Phi}}_j$ and λ_j are the DMD modes and eigenvalues, respectively. In the DMD algorithm, the temporal coefficients $\nu_{j,k} = \lambda_j^{k-1}$ are intrinsically defined by the DMD eigenvalues and the time step k. In principle, (5) can be used to model the flow dynamics outside the time-horizon T of the snapshots, where $T = (N-1)\Delta t$.

2.2. Selection of modes

In the previous section, the classical DMD algorithm was presented as introduced in [1]. So far, we have not discussed the question of modes' selection which is however central in model reduction. Indeed, we are often interested to select a small subset of modes (here DMD modes $\widetilde{\boldsymbol{\Phi}}_j$) that gives a good approximation of a physical quantity of interest. In POD or balanced truncation, the criterion of selection is clear since the modes are by construction ranked by energy level through the POD or Hankel eigenvalues. In DMD, there is no natural way to rank the contributions of the different DMD modes. Moreover, the modes are not orthogonal what further complicated the choice. At this point, we can return to the DMD algorithm for introducing different criteria depending on what can be considered as the "most important" for the physical analysis. One advantage of DMD compared to POD is that each DMD mode j is associated with a pulsation ω_j and a growth rate σ_j. Indeed, by virtue of the discrete-continuous time equivalence, we have $\nu_{j,k} = \nu_j(t_k) = \lambda_j^{t_k/\Delta t}$ with $t_k = (k-1)\Delta t$, and

$$\boldsymbol{v}_k = \sum_{j=1}^{N-1} \lambda_j^{k-1} \widetilde{\boldsymbol{\Phi}}_j = \sum_{j=1}^{N-1} \mathrm{e}^{(\sigma_j + \mathrm{i}\omega_j)t_k} \widetilde{\boldsymbol{\Phi}}_j \quad \text{with } \sigma_j = \frac{\log(|\lambda_j|)}{\Delta t} \text{ and } \omega_j = \frac{\arg(\lambda_j)}{\Delta t} \tag{6}$$

The DMD modes can then be selected based on their amplitude $\|\widetilde{\boldsymbol{\Phi}}_j\|^2$, or based on their frequency/growth rate. The amplitude criterion is not perfect because there exist modes with very high amplitudes but which are very fast damped. Modes' selection based on frequency/growth rate is also not perfect since it relies on a priori physical knowledge. If we know in advance modes that are essential in the flow physics, there is naturally possibility to incorporate them in the model but the process is not fair. To avoid these difficulties, we introduce in the following (see Section 5) a new energetic criterion for which the amplitude of the mode is weighted by its temporal coefficient. For any mode j, this energy is defined as:

$$E_j = \frac{1}{T}\int_0^T \left\| \widetilde{\boldsymbol{\Phi}}_j \lambda_j^{t/\Delta t} \right\|^2 \mathrm{d}t = \|\widetilde{\boldsymbol{\Phi}}_j\|^2 \frac{\mathrm{e}^{2\sigma_j T} - 1}{2\sigma_j T} \tag{7}$$

Note that if we sum E_j for $j = 1, \ldots, N-1$, the flow energy is not obtained since the DMD modes are not orthogonal.

3. Optimized DMD

As discussed in Section 2.2, modes' selection is not trivial in the classical DMD algorithm. Indeed, the non-orthogonality of the DMD modes may raise the projection error while increasing the order of the DMD basis. Moreover, in the classical DMD algorithm, the residual depends only on the last snapshot of the sequence since the companion matrix C is fully defined by the coefficients c_k $(k = 1, \ldots, N-1)$ i.e. by the snapshot $\boldsymbol{v}_N$. As such, the DMD results are more sensitive to variations of $\boldsymbol{v}_N$ (level of noise, for instance) than to variations in other snapshots. To address these issues, we propose to use the *optimized DMD* as recently introduced by Chen et al. [5]. Given $N_o < N$, this method consists in seeking complex scalars $\{\hat{\lambda}_j\}_{j=1}^{N_o}$ and vectors $\{\hat{\boldsymbol{\Phi}}_j\}_{j=1}^{N_o}$ such that

$$\boldsymbol{v}_k = \sum_{j=1}^{N_o} \hat{\boldsymbol{\Phi}}_j \hat{\lambda}_j^{k-1} + \boldsymbol{r_k} \quad k = 1, \ldots, N, \quad \text{and} \quad \Gamma = \sum_{k=1}^{N} \|\boldsymbol{r}_k\|_2^2 \tag{8}$$

is minimized. In optimized DMD, the number of modes that is searched is also a parameter of the method leading by construction to a reduced-order model of desirable size. In the original algorithm presented in Chen et al. [5], the modes were

determined with a global optimization technique combining simulated annealing and the Nelder–Mead simplex method. Here, we improved the original algorithm and determined analytically the gradient of Γ with respect to the variation of the eigenvalues $\hat{\lambda}_j$ [6]. Two advantages of this technique are that all the optimization is done in a space of size N_o and that we can use a descent method for increasing the speed of convergence.

4. DMD-based reduced-order model

In the classical DMD algorithm or in optimized DMD, the assumption of linear dynamics leads jointly to the extraction of a basis for the flow and to the introduction of a time propagator. We will present in this section how to use these informations for deriving a reduced-order model based on DMD.

4.1. DMD time propagator

By definition, DMD identifies the linear operator which represents at best a sequence of snapshots. A direct consequence of the linear assumption is that each DMD mode contains only a single frequency while POD modes, which capture the most energetic structures, give modes that contain several frequencies. The effect of that on the time stepping operator is clearly visible in the reconstruction equation (5). Indeed, the contribution of each DMD mode is weighted by the corresponding eigenvalue raised to the index of the time step. Hence, if the linear assumption corresponds really to the physical phenomena (linear or weakly non-linear systems, data lying on a limit cycle) then it has a sense to propagate the state using this operator. Likewise, if the DMD modes are numerically well determined, especially when data come from experiments, or if the modes used in (5) are well selected, then approximating the future dynamical behaviour of the system in the DMD subspace may be relevant. The validity of the DMD time propagator is then strongly dependent on the data used for determining the DMD modes, and especially on the linearity assumption.

4.2. Petrov–Galerkin projection

In (5), the temporal coefficients of the states are only depending on the DMD eigenvalues and on the time index k. This equation can be considered as a pure kinematic description of the flow since the dynamics is artificially introduced through the linear assumption at the heart of DMD. For increasing the chance to derive a model which can represent the long-term flow dynamics the information that the snapshots are governed by underlying dynamical principles (Navier–Stokes equations in our case) should be incorporated in the modelling step. This is particularly true when the linearity assumption is questionable.

In model reduction, projection methods are very often used for deriving reduced-order models. Starting with a set of trial functions $\boldsymbol{\Phi}_j$, and another set of test functions $\boldsymbol{\Psi}_j$, a weak form of the Navier–Stokes equations is determined. After development, it leads to a quadratic dynamical system given by

$$\sum_{j=1}^{N_{\text{gal}}} G_{ij} \frac{\mathrm{d}v_j}{\mathrm{d}t}(t) = C_i + \sum_{j=1}^{N_{\text{gal}}} L_{ij} v_j(t) + \sum_{j=1}^{N_{\text{gal}}} \sum_{k=1}^{N_{\text{gal}}} Q_{ijk} v_j(t) v_k(t) \tag{9}$$

where N_{gal} is the number of modes kept in the expansion. Different choices are now possible for the test functions. A natural choice corresponds to the adjoint DMD modes since they are biorthonormal to the full DMD basis. The adjoint DMD modes, approximation of the eigenvectors of the Perron–Froebenius operator, are computed by determining the eigenvectors of the adjoint of the companion matrix, and then by multiplying by the pseudo-inverse of the matrix of the snapshots. However, some of these adjoint modes are numerically corrupted due to the ill-conditioning of their determination. For this reason, we have retained in the following the case of the Galerkin projection where $\boldsymbol{\Psi}_j = \boldsymbol{\Phi}_j$, and used for $\boldsymbol{\Phi}_j$ the optimized DMD modes. Since the DMD modes are not orthonormal, G is a full Hermitian matrix, called Gram matrix, and has to be inverted once for integrating in time (9). Finally, the coefficients C_i associated with the pressure term are neglected, assuming that the integral of the pressure around the boundaries of the domain is very small.

4.3. Data assimilation

In the previous subsections, two different approaches were presented for determining a reduced-order model based on DMD. The first method (see Section 4.1) comes directly from the DMD algorithm and as such is purely kinematic while the second method (see Section 4.2) does not incorporate all the dynamical informations coming from DMD. The objective of this section is to combine the two sources of informations for deriving a more representative dynamical system.

Data assimilation [7] is the right framework for combining heterogeneous observations with the underlying dynamical principles governing the system under observation to estimate at best physical quantities. Here, we apply the four-dimensional variational approach of data assimilation (4D-Var). More precisely, we seek for the initial condition perturbation $\boldsymbol{\eta}$ and the model coefficients $\boldsymbol{c} = \{C_i, L_{ij}, Q_{ijk}\}$ of (9) such that the solutions of the dynamical model tend to the model coefficients obtained directly by optimized DMD. For improving the numerical convergence of the optimization problem,

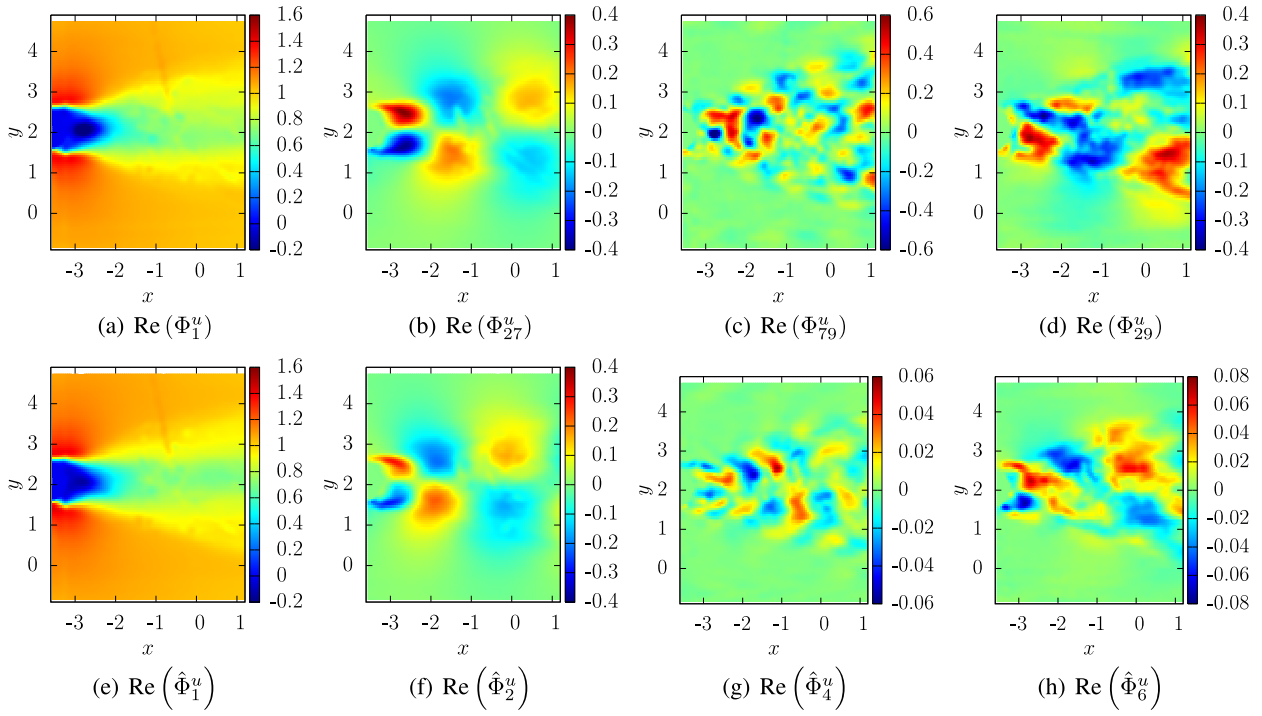

Fig. 1. (Colour online.) Comparison of modes obtained by classical DMD ($\boldsymbol{\Phi}_i$) and optimized DMD ($\hat{\boldsymbol{\Phi}}_i$). For reasons of space limitation, only the streamwise components are plotted.

background errors which penalize the variation between the background states $(0, \boldsymbol{c}^b)$ and the estimated values are introduced. The corresponding cost functional reads:

$$J(\boldsymbol{\eta}, \boldsymbol{c}) = \frac{1}{2} \sum_{k=1}^{N_t} \sum_{j=1}^{N_o} \left(\nu_j(t_k; \boldsymbol{\eta}, \boldsymbol{c}) - \hat{\lambda}_j^{k-1}\right)^2 + \frac{\sigma_\eta}{2} \|\boldsymbol{\eta}\|^2 + \frac{\sigma_c}{2} \left\|\boldsymbol{c} - \boldsymbol{c}^b\right\|^2 \tag{10}$$

where N_t are the number of time steps for the time horizon of the reduced-order model, and where σ_η and σ_c are penalization terms which give more or less weight in the background solutions. The background states are found by application of the Galerkin projection onto the optimized DMD modes.

5. Results

The classical and optimized DMD algorithms have first been tested on numerical data corresponding to a cylinder wake at $Re = 200$ [6]. In that case, the optimized DMD modes correspond to the most energetic classical DMD modes as selected according to criterion (7). These modes distinguish from the POD modes only by the presence of single frequencies. Moreover, reduced-order models have been derived successfully either by the DMD time propagator, or through the Galerkin projection onto DMD modes selected by the energetic criterion E_j.

In this paper, 2D–2C PIV data are considered for a cylinder wake at $Re = 13\,000$. The database contains $N_s = 1000$ snapshots sampled at the frequency $f_s = 1$ kHz. First, the classical DMD algorithm was applied for $N = N_s$. In the case of experimental data, modes' selection becomes extremely hard, and considering too many DMD modes in the reconstruction of snapshots may lead to high level of errors due to the non-orthogonality of the basis. Optimized DMD has then been performed on the first 256 snapshots ($T = 30$). The optimization problem linked to the optimized DMD is solved through a gradient descent algorithm. The initial conditions are DMD modes selected with the energetic criterion E_j. In Fig. 1, we compare the initial conditions of the optimized DMD algorithm (top row) and the optimized DMD modes determined at convergence (bottom row). For the two first optimized DMD modes, only small changes have been made compared to the DMD modes used for the initialization. For the higher order optimized DMD modes, the changes are much more important. Then, we reconstructed the original snapshots from 7 modes obtained by classical DMD (modes' selection based on E_j) and optimized DMD (see Fig. 2 for the fifth snapshot). As expected, the L^2-norm error of reconstruction is lower for the optimized DMD than for the classical algorithm. For the optimized DMD, we obtain very good filtered approximation of the original snapshots.

Finally, a 4D-Var approach has been performed in the optimized DMD subspace for $\sigma_\eta = \sigma_c = 1$. This assimilation was made in a time window of size $T = 30$ and then, the optimal solution was forecast over a time length equal to $3T$. In Fig. 3, we compare the temporal coefficients ν_j obtained by optimized DMD and 4D-Var to the projection of the snapshots on the optimized DMD modes. For the long-term horizon, the 4D-Var solution outperforms the optimized DMD.

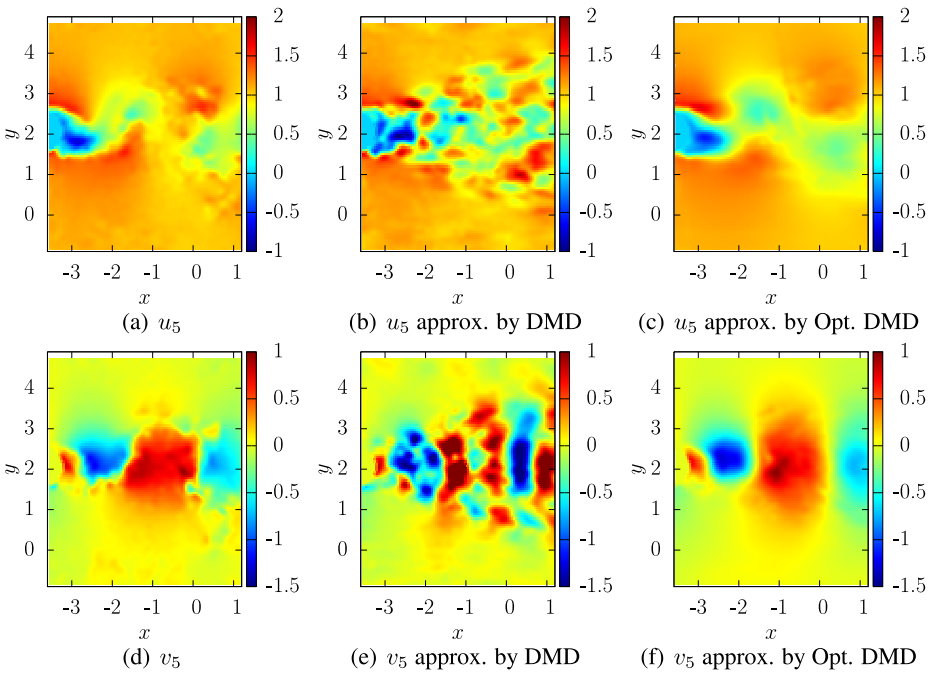

Fig. 2. (Colour online.) Reconstruction of the snapshot $\boldsymbol{v}_5 = (u_5, v_5)$ with *seven* modes obtained by classical DMD and optimized DMD. The corresponding L^2-norm errors are 45.6% (classical DMD) and 15.6% (optimized DMD). The selection of the DMD modes used in the reconstruction is based on the energetic criterion given by (7).

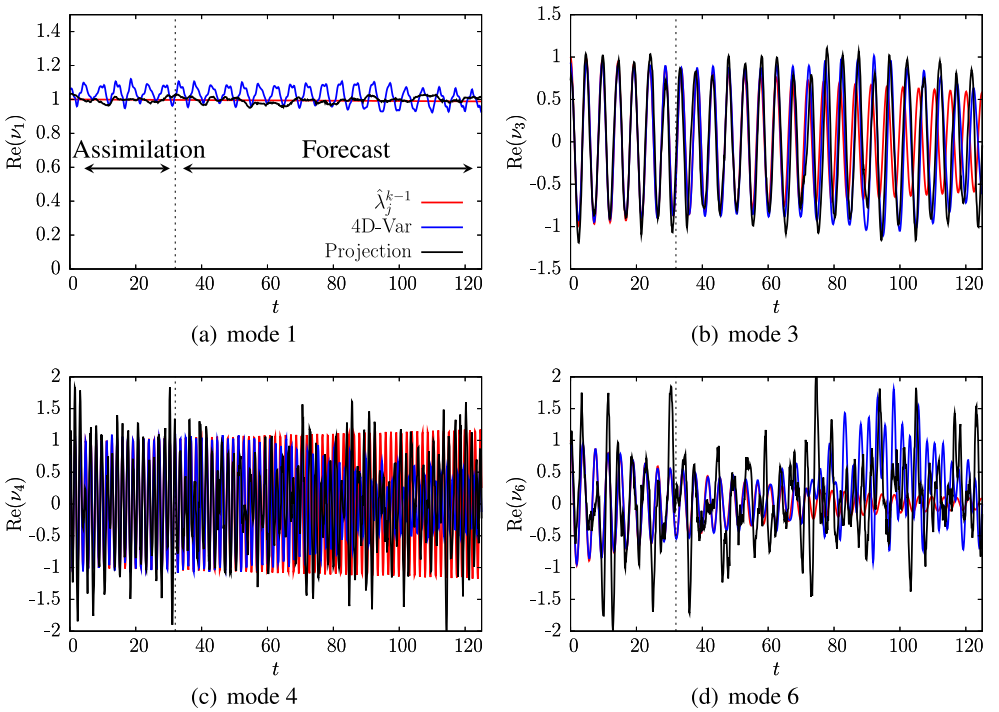

Fig. 3. (Colour online.) Comparison of temporal coefficients v_j (real parts) obtained by optimized DMD, 4D-Var ($\sigma_\eta = \sigma_c = 1$) and projection on the optimized DMD modes. 4D-Var is solved in the optimized DMD space over the assimilation window. The optimal solution is then forecast.

6. Conclusion

The classical and optimized DMD algorithms were applied in the paper with the objective to derive reduced-order models. In the classical DMD algorithm, modes' selection may become difficult since most of the time we do not know if we should privilege mode's energy, frequency behaviour or growth rate. For this reason, we introduce a new energetic criterion that incorporates the growth rate of the modes. An alternative is to use directly the optimized DMD algorithm as proposed recently by Chen et al. [5] since the number of modes is also a parameter of the method. In the DMD framework, the time propagation is included in the approach. Hence, we have guaranty that the temporal evolution of the system can be well reproduced over the time window of the snapshots where the linear approximation should hold. To improve this result, we first derive a reduced-order model obtained by Galerkin projection of the Navier–Stokes equations onto the DMD modes. We then introduce a 4D-Var approach to combine the informations coming from the DMD, and those coming from the Galerkin projection. Finally, we showed that for the long-term horizon, the 4D-Var solution outperforms the optimized DMD.

References

[1] P.J. Schmid, Dynamic mode decomposition of numerical and experimental data, J. Fluid Mech. 656 (7) (2010) 5–28.
[2] C.W. Rowley, I. Mezić, S. Bagheri, P. Schlatter, D.S. Henningson, Spectral analysis of nonlinear flows, J. Fluid Mech. 641 (2009) 115–127.
[3] I. Mezić, Spectral properties of dynamical systems, model reduction and decompositions, Nonlinear Dyn. 41 (1–3) (2005) 309–325.
[4] S. Bagheri, Analysis and control of transitional shear flows using global modes, PhD thesis, Department of Mechanics, Royal Institute of Technology, SE-100 44, Stockholm, Sweden, 2010.
[5] K.K. Chen, J.H. Tu, C.W. Rowley, Variants of dynamic mode decomposition: boundary condition, Koopman, and Fourier analyses, J. Nonlinear Sci. 22 (6) (2012) 887–915.
[6] G. Tissot, Réduction de modèle et contrôle d'écoulements, PhD thesis, Université de Poitiers, 2013.
[7] L. Cordier, B.R. Noack, G. Daviller, G. Tissot, G. Lehnasch, J. Delville, M. Balajewicz, R. Niven, Control-oriented model identification strategy, Exp. Fluids 54 (2013) 1–21 (special issue "Flow Control").