A Dynamic Mode Decomposition approach for reduced-order Data Assimilation in the presence of parameter uncertainties

Arthur Vervynck[★][†], *Vincent Mons*[★] *and Frédéric Champagnat*[★★]
[★] *DAAA, ONERA, Université Paris Saclay, F-92190 Meudon, France*
[★★] *DTIS, ONERA, Université Paris Saclay, F-91123 Palaiseau, France*
arthur.vervynck@onera.fr · vincent.mons@onera.fr · frederic.champagnat@onera.fr
[†]Corresponding author

Abstract

Data assimilation may be expensive, often requiring numerous calls to a numerical solver. We here consider replacing the numerical model with a reduced-order one based on the Dynamic Mode Decomposition (DMD). To tackle parameter uncertainties, it is proposed to design the reduced-order model from a catalogue of DMDs corresponding to different values of the uncertain parameters. The so-obtained dynamical model is then used in a Kalman filter for data assimilation. The proposed methodology is successfully applied to the cost-efficient estimation of the flow around a circular cylinder from sparse measurements, considering the Reynolds number of the flow as badly known.

1. Introduction

Measurement techniques such as PIV (Particle Image Velocimetry) or PTV (Particle Tracking Velocimetry) may provide relatively detailed information about flows of interest. However, the spatial and temporal resolutions of experimental data remain limited, not to mention measurement errors. This may motivate the use of data assimilation[2] techniques, which enable the use of Computational Fluid Dynamics (CFD) models to provide a full flow estimation from measurements.[4]

Data assimilation methods can be separated into two categories: variational and sequential methods. Concerning variational methods, a cost function is minimized through a gradient-based descent method, relying on the adjoint of CFD model. Reaching convergence typically requires numerous calls to the direct and adjoint CFD solvers. In sequential methods such as the Kalman filter, the covariance of the state has to be evaluated and advanced in time, which may be costly if the dimension of the (discretized) state is large, which is generally the case in CFD applications. In both cases, data assimilation is thus a numerically expensive procedure. Still, Kalman filters have the advantage of being relatively non-intrusive compared to variational approaches and are particularly suited for state estimation. The aim of the present study is thus to develop a Kalman filter-based methodology with low computational cost to significantly foster the application of data assimilation to flow reconstruction from limited measurements.

The most expensive step in the Kalman filter is the propagation step, i.e. the advancement in time, usually based on a CFD model, of the state estimation and of its statistics between two measurement instants. Accordingly, a strategy to reduce the computational cost of Kalman filtering may consist in replacing the CFD model with a reduced-order one. This may be achieved through, among others, the Dynamic mode Decomposition (DMD) proposed by Schmid,[10] which provides the best linear approximation of a dynamical model from an ensemble of snapshots. The computational cost of advancing in time the flow based on DMD is almost negligible compared to the use of a CFD model. The combination of DMD with Kalman filtering has been investigated by Fathi et al.,[3] Tsolovikos et al.[12] and Iungo et al.,[6] among others. DMD has also been considered in conjunction with variational data assimilation by Tissot et al.[11]

In order to ensure that the DMD model correctly captures the dynamics of the flow to reconstruct, a necessary condition is that the CFD model that we want to approximate through DMD should itself be able to correctly describe the flow. Aside from numerical and modeling errors, which are not considered here, this requires in particular that the snapshots that are generated by the CFD model to form the DMD should correspond to the same physical parameters as for the flow to reconstruct. Otherwise, the estimated flow would be restricted in a subspace that does not include the true one. However, when considering experimental data, the corresponding parameters such as the angle of attack, the Mach number or the Reynolds number, among others, are not always known with

high accuracy. Such parameter uncertainties should thus be taken into account when building the DMD-based reduced-order model.

The main proposal of this study is to address these parameter uncertainties based on a catalogue of DMDs that correspond to different values of the badly-known parameters. While the use of several DMDs for data assimilation has recently been investigated by Tsolovikos et al.,[12] a key feature of the present approach is the generation of a single reduced-order model from the catalogue of DMDs. The reconstructed flow thus belongs to the subspace spanned by the union of the DMDs, thus favoring the fidelity of the reconstructed flow. As data assimilation will be here performed through Kalman filtering in the reduced space defined by the catalogue of DMDs, the proposed approach will be referred to as the Catalogue Kalman Filter (CKF), or as the Catalogue Ensemble Kalman Filter (CEnKF) when employing the EnKF.[2]

In this paper, the CKF methodology is exposed and assessed for unsteady flow reconstruction from sparse measurements. The paper is organized as follows. The Dynamic Mode Decomposition and the Kalman Filter approaches, along with their combination to perform reduced-order data assimilation, are first presented in section 2. The flow configuration that is considered to apply reduced-order data assimilation, namely the laminar flow around a circular cylinder at the Reynolds number Re = 100, is detailed in section 3. First reconstruction results from pointwise measurements are reported in section 4, where it is shown how inaccuracy in the Reynolds number leads to deteriorated results. This motivates the proposed CKF methodology, which is described in section 5. The CKF scheme is then applied in section 6, allowing to obtain better reconstruction results compared to the use of any single DMD in the catalogue. Conclusions and perspectives are drawn in section 7.

2. Kalman filtering based on the Dynamic Mode Decomposition

2.1 Kalman Filter

To perform data assimilation, and here focusing in unsteady state estimation, we here consider the use of the Kalman Filter[7] and of its variants. The Kalman Filter methodology is decomposed into two steps: forecast (superscript ^{f}) and analysis (superscript ^{a}), which are detailed in the following for a linear dynamical model. We define the following dynamical and observational models for a state $\boldsymbol{x}_k$ and measurements $\boldsymbol{y}_k$, where the subscript k refers to time:

$$\boldsymbol{x}_k = \mathbf{M}\boldsymbol{x}_{k-1} + \boldsymbol{\eta}_k \tag{1a}$$

$$\boldsymbol{y}_k = \mathbf{H}\boldsymbol{x}_k + \boldsymbol{\epsilon}_k \tag{1b}$$

where $\boldsymbol{\eta}_k$ and $\boldsymbol{\epsilon}_k$ are Gaussian noises with covariance matrices $\mathbf{Q}_k$ and $\mathbf{R}_k$ respectively, $\mathbf{M}$ is a linear dynamic model and $\mathbf{H}$ the operator that maps the state space to the observation space. The aim of data assimilation is thus to infer the full state $\boldsymbol{x}_k$ from the measurements $\boldsymbol{y}_k$.

During the forecast step, the estimated state as well as the associated covariance matrix $\mathbf{P}_k^{f,a}$ are propagated to the next time-step k according to:

$$\boldsymbol{x}_k^f = \mathbf{M}\boldsymbol{x}_{k-1}^a \tag{2a}$$

$$\mathbf{P}_k^f = \mathbf{M}\mathbf{P}_{k-1}^a\mathbf{M}^{\mathrm{T}} + \mathbf{Q}_{k-1} \tag{2b}$$

In the analysis step, the estimated state and its covariance matrix are corrected using the observations $\boldsymbol{y}_k$ following:

$$\boldsymbol{x}_k^a = \boldsymbol{x}_k^f + \mathbf{K}_k\left(\boldsymbol{y}_k - \mathbf{H}\boldsymbol{x}_k^f\right) \tag{3a}$$

$$\mathbf{P}_k^a = (\mathbf{I}_k - \mathbf{K}_k\mathbf{H})\,\mathbf{P}_k^f \tag{3b}$$

where $\mathbf{K}_k$ is the Kalman gain defined by:

$$\mathbf{K}_k = \mathbf{P}_k^f\mathbf{H}^{\mathrm{T}}\left(\mathbf{H}\mathbf{P}_k^f\mathbf{H}^{\mathrm{T}} + \mathbf{R}_k\right)^{-1} \tag{4}$$

In order to handle nonlinear dynamics and high-dimensional systems, on may instead rely on the Ensemble Kalman Filter (EnKF).[2] The difference between this filter and the standard Kalman Filter is that the covariance matrix for the estimated state is not explicitly formed but is approximated and advanced in time based on an ensemble of N states:

$$\mathbf{E}^{f,a} = \left[\boldsymbol{x}_{(1)}^{f,a},...,\boldsymbol{x}_{(\mathrm{N})}^{f,a}\right] \tag{5}$$

Even for the EnKF, the forecast step remains computationally demanding in the case of CFD, as it requires performing N simulations for the forecast step, with $N \sim O(10) - O(100)$. In the next subsections, a reduced-order modeling approach is described to alleviate the cost of Kalman filtering.

2.2 Dynamic Mode Decomposition

The Dynamic Mode Decomposition (DMD) was proposed by Schmid[10] as a data-driven method to extract dynamically relevant features in flow fields. The algorithm that is here used to compute the DMD is referred as the "exact" DMD.[13] We consider a set of m pairs of consecutive snapshots $\{\boldsymbol{x}_k, \boldsymbol{x}_{k+1}\}_{k=1}^m$. We then define $\mathbf{X}$ and $\mathbf{X}'$ as the matrices containing all first and second elements of the m snapshot pairs, respectively:

$$\mathbf{X} = [\boldsymbol{x}_1 \ \boldsymbol{x}_2 \ \cdots \ \boldsymbol{x}_m] \tag{6a}$$

$$\mathbf{X}' = [\boldsymbol{x}_2 \ \boldsymbol{x}_3 \ \cdots \ \boldsymbol{x}_{m+1}] \tag{6b}$$

DMD consists in determining the matrix $\mathbf{A}$ that gives the best linear approximation $\mathbf{X}' \approx \mathbf{A}\mathbf{X}$. If the state is of high dimension, $\mathbf{A}$ can first be projected onto the leading r Proper Orthogonal Decomposition (POD) modes $\mathbf{U}_r$ as detailed by Tu et al.[13] This is performed through the Singular Value Decomposition (SVD) of $\mathbf{X}$: $\mathbf{X} = \mathbf{U}\Sigma\mathbf{V}^*$, where * denotes the complex conjugate transpose and $\mathbf{U}$ is the matrix containing the POD modes. It is then possible to truncate the rank of the matrix $\mathbf{X}$ to a value r by selecting the first r columns and/or rows of $\mathbf{U}$, Σ and $\mathbf{V}$: $\mathbf{X} \approx \mathbf{U}_r\Sigma_r\mathbf{V}_r^*$. The matrix $\mathbf{A}$ may therefore be evaluated according to:

$$\mathbf{A} = \mathbf{X}'\mathbf{X}^\dagger = \mathbf{X}'\mathbf{V}\Sigma^{-1}\mathbf{U}^* \approx \mathbf{X}'\mathbf{V}_r\Sigma_r^{-1}\mathbf{U}_r^* \tag{7}$$

where † denotes the Moore-Penrose pseudo inverse of a matrix. $\mathbf{A}$ may thus be used to advance the state $\boldsymbol{x}_k$ in time according to:

$$\boldsymbol{x}_{k+1} = \mathbf{A}\boldsymbol{x}_k \tag{8}$$

The DMD modes, which are denoted by Φ, correspond to the r eigenvectors of $\mathbf{A}$. They thus verify:

$$\mathbf{A}\Phi = \Phi\Lambda \tag{9}$$

with Λ the diagonal matrix of DMD eigenvalues. Based on this eigendecomposition, the state $\boldsymbol{x}_k$ may be expressed as the linear combination of the DMD modes, each oscillating at a single frequency as prescribed by the phase of the DMD eigenvalues, according to:

$$\boldsymbol{x}_k = \sum_{j=1}^r \alpha_j \boldsymbol{\phi}_j \lambda_j^{k-1} \tag{10}$$

with $\boldsymbol{\phi}_j$, λ_j and α_j the jth DMD mode, DMD eigenvalue and the projection of the initial condition on the mode $\boldsymbol{\phi}_j$ ($\boldsymbol{x}_0 = \sum_{j=1}^r \alpha_j \boldsymbol{\phi}_j$) respectively. Contrary to the POD modes which are straightforwardly ordered by decreasing energy, there exists different ways to arrange DMD modes. In this paper, the modes are ordered as described by Kou et al.[8] to take into account the temporal evolution of their amplitude until the time t_K based on the quantity I_j for the mode $\boldsymbol{\phi}_j$:

$$I_j = \sum_{i=1}^K |\alpha_j \lambda_j^{i-1}| \|\phi_j\| \delta t \tag{11}$$

The DMD modes calculated are composed of complex conjugate pairs apart from the first mode, which corresponds to the mean flow and is real and of null frequency. In other words, the mean flow is here not subtracted from the snapshots when computing the DMD[1] in order to allow for the possibility of updating the latter through data assimilation.

2.3 Reduced-order Kalman filtering

The Kalman Filter and the DMD can be combined in order to form a reduced-order filter which is drastically less expensive compared to the use of a CFD model for the forecast step. Considering the following reduced state based on the DMD modes:

$$\tilde{\boldsymbol{x}}_k = \Phi^\dagger \boldsymbol{x}_k \tag{12}$$

its advancement in time is performed through:

$$\tilde{\boldsymbol{x}}_{k+1} = \Lambda \tilde{\boldsymbol{x}}_k \tag{13}$$

A DMD APPROACH FOR DATA ASSIMILATION

which is easily evaluated as it corresponds to a diagonal linear dynamical model of low order, i.e. of size r from the truncated SVD step in section 2.2. Therefore, the forecast step in the Kalman filter for the state as given by (13) becomes virtually inexpensive compared to the use of a CFD model. While the standard Kalman Filter or the EnKF could be used indifferently for such a reduced state with linear dynamics, we will here rely on the EnKF. In other words, the covariance matrix of the reduced state will be advanced in time through N evaluations of (13), with N the number of ensemble numbers (see section 2.1).

It may be noted that the relationship between observations and the reduced state may be written as:

$$\boldsymbol{y}_k = \mathbf{H}\Phi\tilde{\boldsymbol{x}}_k = \widetilde{\mathbf{H}}\tilde{\boldsymbol{x}}_k \tag{14}$$

Accordingly, the analysis step in reduced space is similar to equation (3), replacing $\boldsymbol{x}_k^{f,a}$ and $\mathbf{H}$ with $\tilde{\boldsymbol{x}}_k^{f,a}$ and $\widetilde{\mathbf{H}}$, respectively. Such a reduced-order Kalman filtering approach is applied in next sections and will form the basis of the proposed Catalogue (Ensemble) Kalman Filter.

3. Test case and implementation details

(a)   (b)
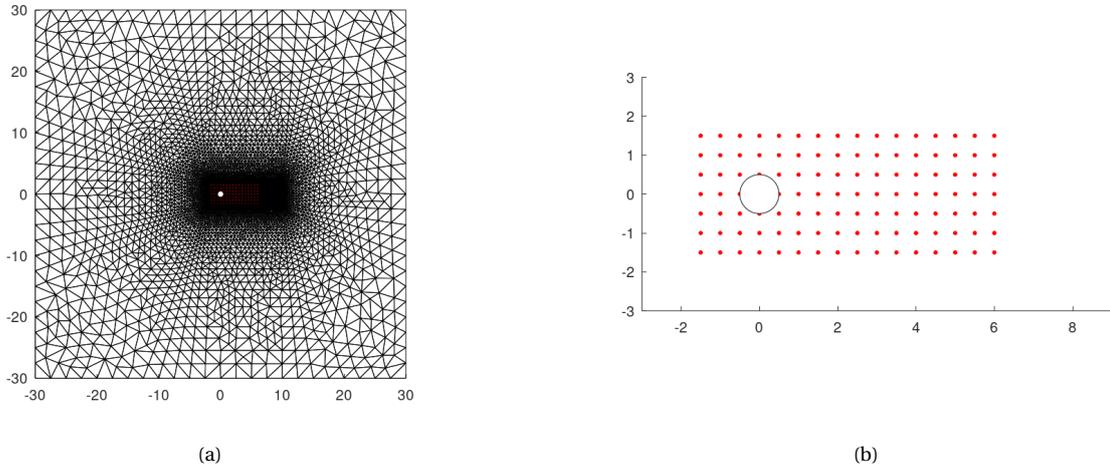
Figure 1: Setup for the 2D flow around a cylinder: (a) computational domain and mesh, (b) zoom on the cylinder and its wake, the locations of velocity measurements are indicated as red dots.

The above methodologies are applied to the estimation of the flow around a 2D circular cylinder from point-wise measurements. The reference flow, which may be referred to as "truth" in the following, first corresponds to the Reynolds number Re = 100 based on the cylinder diameter and uniform inlet velocity, which are used to nondimensionalize all results in the following. This reference flow is generated numerically through solving the incompressible Navier-Stokes equations relying on the finite element solver FreeFem++.[5] Pointwise synthetic velocity measurements with a spacing of 0.5 (i.e. one diameter radius) in both streamwise and crosstream directions are extracted from this simulation and are perturbed using a white noise with variance $1e-4$. This corresponds to a standard deviation of 1%, which may be representative of typical errors in PIV and PTV data. The mesh and the observation points are illustrated in figure 1. The number of degrees of freedom in the simulation is around $5 \cdot 10^4$, including P2 elements for the two velocity components and P1 elements for pressure.

Considering snapshots that correspond to the true Reynolds number or to other values, DMDs are calculated by taking 400 of them separated each by $\delta t = 0.1$, which leads to over 50 snapshots per period of vortex shedding, as the Strouhal number of the flow is approximately 0.16 at Reynolds numbers around 100. As detailed previously, the DMD modes can be calculated by projecting the snapshot matrix on the r dominant POD modes. The energy of a POD mode is obtained by taking the square of the associated singular value σ^2. Thus, the cumulative energy up to the j-th POD mode out of a total of r modes is:

$$\mathrm{E} = \frac{\sum_{i=1}^{j} \sigma_i^2}{\sum_{i=1}^{r} \sigma_i^2} \tag{15}$$

The energy left after selecting the first j POD modes is then $\mathrm{E}_{left} = 1 - \mathrm{E} = 1 - \sum_{i=1}^{j} \sigma_i^2 / \sum_{i=1}^{r} \sigma_i^2$. This quantity is represented on figure 2 for Re = 100. This shows that by selecting r = 19 modes, almost no energy is left and the
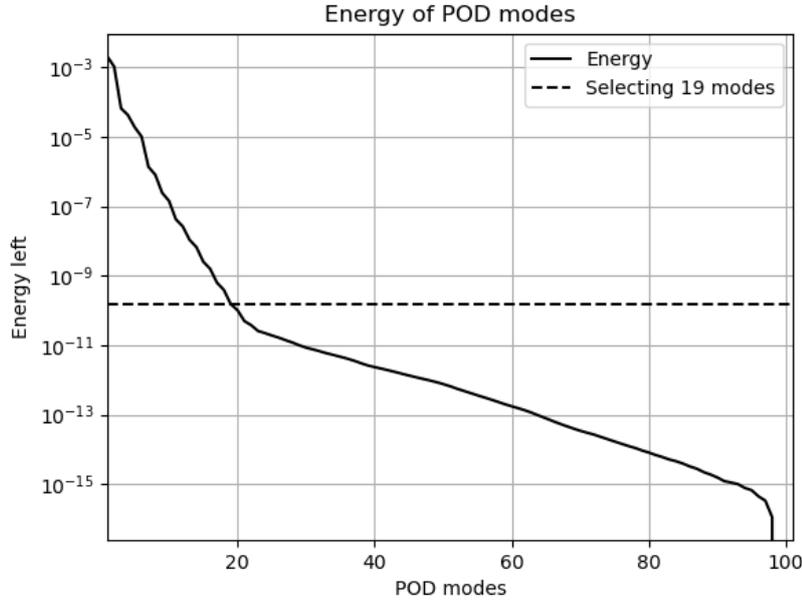
Figure 2: Energy left versus the number of retained POD modes.

remaining POD modes have negligible energies. On the POD basis, r DMD modes are calculated, which thus leads in the present case to 19 DMD modes per simulation. An odd number of DMD modes per Reynolds number allows to have pairs of complex conjugate modes and the mean flow (which corresponds to the first DMD mode). The real part of the voriticity field for some these modes is illustrated in figure 3.

(a)

(b)

(c)
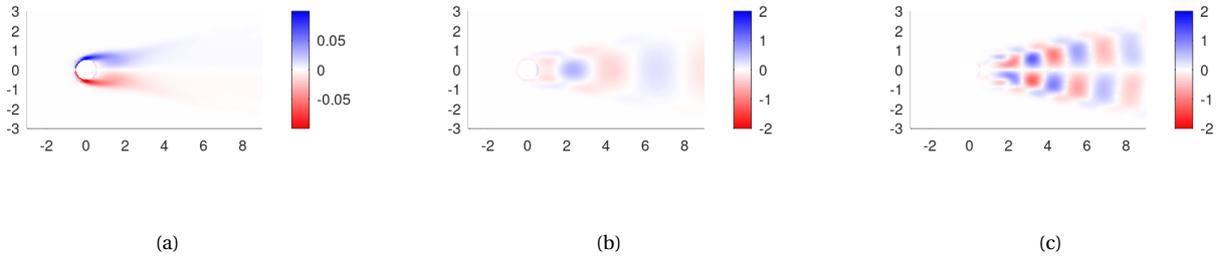
Figure 3: Real part of the vorticity field for the (a) first (mean flow), (b) second and (c) fourth DMD modes.

Observations are available every $\delta t = 0.1$, which matches the time step of the DMD. In order to generate the initial ensemble for the EnKF, the statistics (mean and covariance matrix) of the reduced state $\tilde{x}_k$ in (12) are first straightforwardly calculated from the snapshots. Then N = 19 random samples are drawn using a Gaussian distribution. The mean initial condition should thus correspond to the mean flow, up to finite ensemble size effects. This can be seen in figure 4d, which is discussed in next section. The EnKF is performed using the Python library DAPPER Version 0.9.6[9] during a total time of $t = 150$ to include approximately 25 vortex Shedding periods and several multiplicative inflation parameters (which increase the ensemble variance) were tested. As the reduced state $\tilde{x}_k$ is complex, real and imaginary parts are distinguished in order to allow the use of a real filter.

4. Application of reduced-order Kalman filtering

4.1 Using a DMD at the true Reynolds number

Reduced-order Kalman filtering (see section 2.3) is here first applied when the DMD is built from snapshots corresponding to the true Reynolds number, i.e. Re = 100. Reconstruction results are illustrated in figure 4. The first row

of this figure reports the vorticity field of the true flow to reconstruct at three different times. The estimated flow from reduced-order Kalman filtering is illustrated in the second row, while the third row shows the discrepancies between the true and estimated flows. It appears that, starting from a non-informed initial condition, the reduced-order EnKF successfully estimates the true flow from the pointwise measurements, remaining discrepancies being negligible.

(a) $\boldsymbol{\omega}^t$ for Re 100 at t=0

(b) $\boldsymbol{\omega}^t$ for Re 100 at t=75

(c) $\boldsymbol{\omega}^t$ for Re 100 at t=150

(d) $\boldsymbol{\omega}^{\mathrm{DMD100}}$ at t=0

(e) $\boldsymbol{\omega}^{\mathrm{DMD100}}$ at t=75

(f) $\boldsymbol{\omega}^{\mathrm{DMD100}}$ at t=150

(g) $\boldsymbol{\omega}^{\mathrm{DMD100}} - \boldsymbol{\omega}^t$ at t=0

(h) $\boldsymbol{\omega}^{\mathrm{DMD100}} - \boldsymbol{\omega}^t$ at t=75

(i) $\boldsymbol{\omega}^{\mathrm{DMD100}} - \boldsymbol{\omega}^t$ at t=150
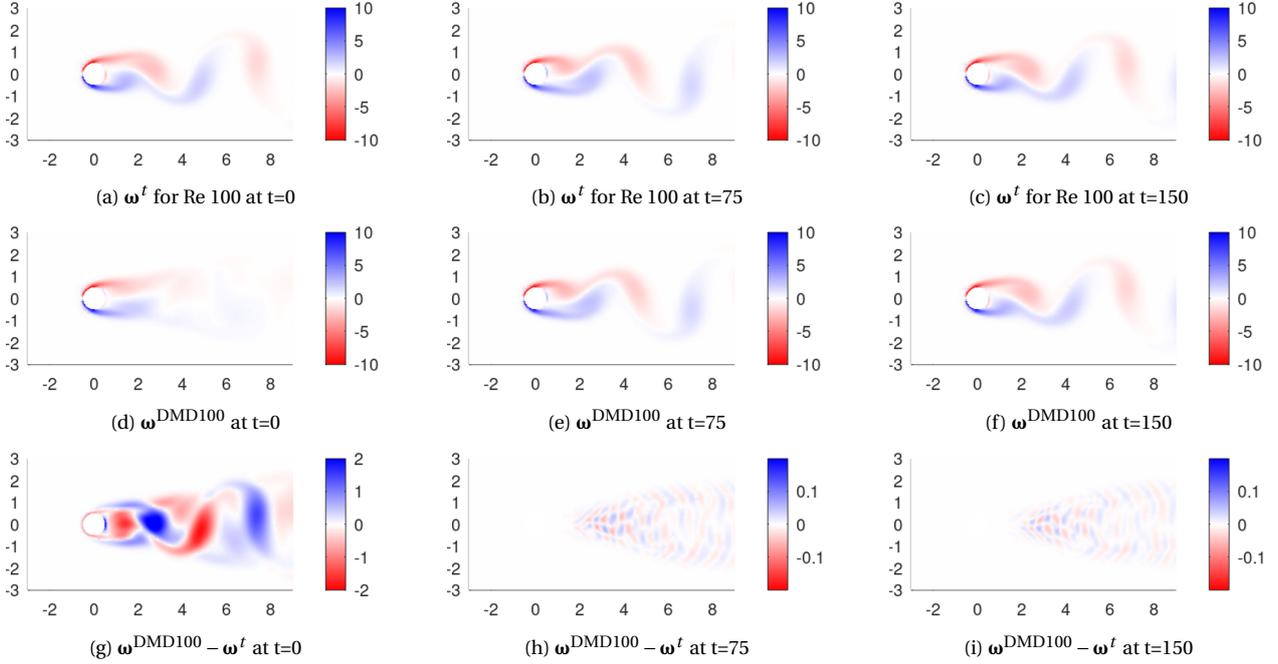
Figure 4: First to third row: vorticity field of the true flow at Re = 100 (first row), estimated vorticity field with reduced-order Kalman filtering based on a DMD at Re = 100 (second row), and difference between estimated and true vorticity fields. First to third column correspond to times $t = 0$, $t = 75$ and $t = 150$.

In order to quantify the quality of the reconstruction, we consider the following instantaneous relative L^2 error between the ensemble mean of the estimated state $\widehat{\boldsymbol{x}}_k$ and the true flow $\boldsymbol{x}_k^t$:

$$\mathrm{E}_k = \frac{\left\| \widehat{\boldsymbol{x}}_k - \boldsymbol{x}_k^t \right\|}{\left\| \boldsymbol{x}_k^t \right\|} \tag{16}$$

The temporal evolution of this error is reported with the full line in figure 5 for the present reconstructed flow. E_k is initially relatively large due to the crude first guess for the initial condition. After a fast decrease, E_k converges towards $\sim 10^{-4}$, confirming the quality of the estimated field. The remaining discrepancies are mainly due to the errors in the synthetic measurements. This is confirmed by a supplementary data assimilation experiment where the variance of the measurement noise is lowered to 10^{-12} (instead of 10^{-4} in the rest of this study), which corresponds to the dashed line in figure 5. E_k is decreased by almost one order of magnitude. For the sake of comparison, the evolution of E_k when simply advancing in time the true initial condition with the DMD model (without data assimilation) is reported with the dotted line. In this case, the value of E_k is thus solely determined by the error due to the DMD approximation, and sets the lowest error level that may be reached by reduced-order Kalman filtering. It appears that the estimated flow when measurement errors have a variance of 10^{-12} reaches this lowest level.

These results confirm the efficacy of reduced-order Kalman filtering when the latter is based on a DMD that has been built from snapshots at the same Reynolds number as for the true flow.

4.2 Inaccuracy on the Reynolds number

We now investigate the case where the Reynolds number of the true flow is badly known. More specifically, it is assumed that the latter is known within a range of $\pm 2\%$ of the true value. Two DMDs are built from snapshots corresponding to Re = 98 and Re = 102. Two data assimilation procedures are then performed where the reduced-order EnKF relies on one of these two DMDs, the corresponding reconstruction results being reported in figure 6.
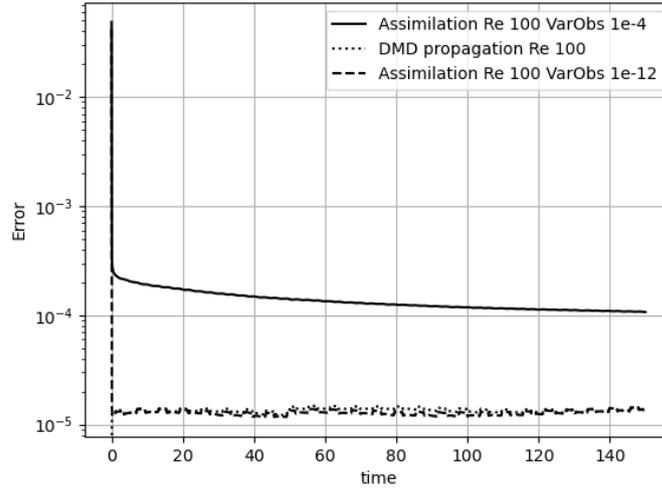
Figure 5: Temporal evolution of the error E_k for the reconstructed flow when the variance of the measurement noise is 10^{-4} (full line) or 10^{-12} (dashed line). E_k is also reported when advancing the true initial condition with the DMD at R$e = 100$ (dotted line).

Compared to the previous case where the reduced-order EnKF is based on a DMD at the true Reynolds number (see third row of figure 4), it appears that the errors in the estimated vorticity fields are significantly higher.

(a) $\omega^{\text{DMD98}} - \omega^t$ at t=0

(b) $\omega^{\text{DMD98}} - \omega^t$ at t=75

(c) $\omega^{\text{DMD98}} - \omega^t$ at t=150

(d) $\omega^{\text{DMD102}} - \omega^t$ at t=0

(e) $\omega^{\text{DMD102}} - \omega^t$ at t=75

(f) $\omega^{\text{DMD102}} - \omega^t$ at t=150
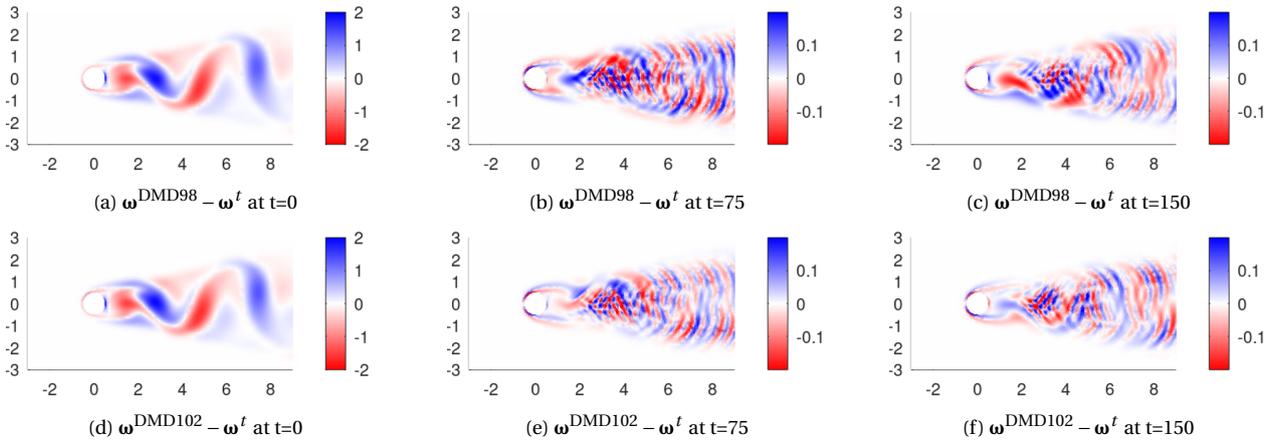
Figure 6: First and second rows: difference between estimated and true vorticity fields when the reduced-order EnKF is based on a DMD at R$e = 98$ (first row) or R$e = 102$ (second row). First to third column correspond to times $t = 0$, $t = 75$ and $t = 150$.

This is confirmed by figure 7 where is reported the temporal evolution of the error E_k for these two cases (dashed and dotted lines). The case where the reduced-order EnKF is based on a DMD at the true Reynolds number is also reported for the sake of comparison (full line). By making an error of $\pm 2\%$ on the Reynolds number, it thus appears that the reconstruction error is increased by more than one order of magnitude. Still, the data assimilation procedure allows to correctly phase the estimated flow and to stabilize the error level. For the sake of comparison, the propagation of the true initial condition by the DMDs at R$e = 98$ and R$e = 102$ are reported with dash-dotted lines, showing a monotonic increase of the errors and confirming the benefits of data assimilation.

To further analyse both model errors that are introduced when building a DMD on snapshots at a wrong Reynolds number and the improvement that may be provided by data assimilation, figure 8a first represents the relative error in the frequency of DMD modes compared to the true flow's ones. These errors are reported for the DMDs at R$e = 98$ and R$e = 102$ (full and dashed lines), but also for DMDs of the reconstructed flows of this section (i.e. for which the DMDs at R$e = 98$ and R$e = 102$ are used in the reduced-order EnKF). Figure 8a confirms that

7
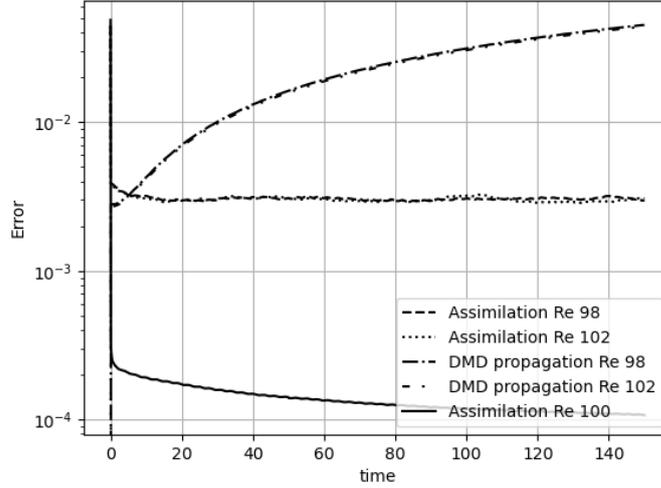
Figure 7: Temporal evolution of the error E_k for the reconstructed flow when the reduce-order Kalman filter is based on a DMD at Re = 98 (dashed line), Re = 102 (dotted line) or at the true Reynolds number Re = 100 (full line). E_k is also reported when advancing the true initial condition with the DMDs at Re = 98 and Re = 102 (dash-dotted lines).
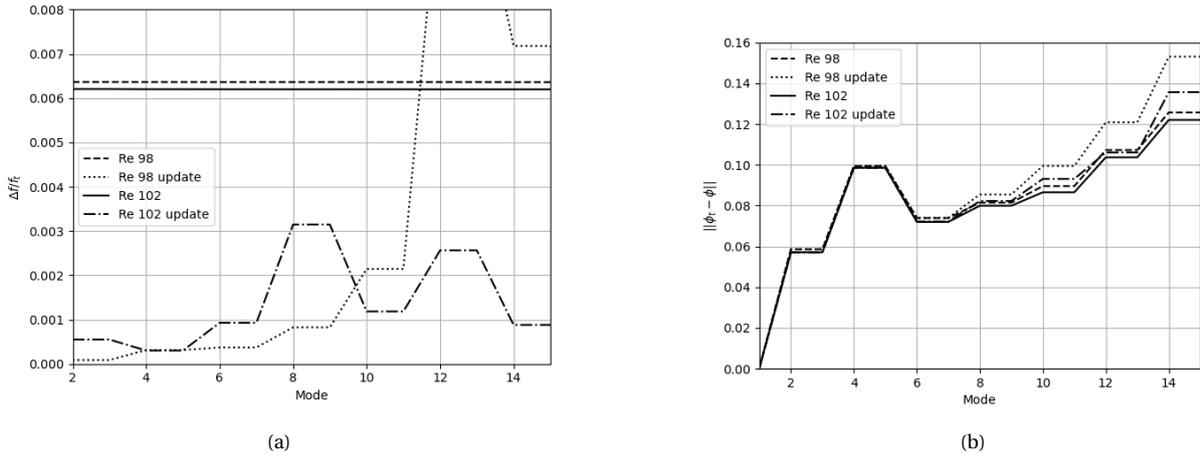
Figure 8: Error on (a) the frequency of DMD modes and (b) on the modes themselves for DMDs at Re = 98 and Re = 102 (dashed and full lines) and for the reconstructed flows that haven been obtained through reduced-order Kalman filtering based on these former DMDs (dotted and dash-dotted lines).

the data assimilation procedure enables a large improvement in the estimation of the frequencies in the true flow, which are almost perfectly recovered. On the other hand, it appears from figure 8b, which reports the norm of the difference between estimated DMD modes and the true ones, that the modes themselves are not corrected throuh data assimilation. Accordingly, only the dynamics of the modes are corrected but not their spatial structures. This is further discussed in next section.

5. Catalogue (Ensemble) Kalman Filter

Results of section 4.2 indicate that the spatial structure of the DMD modes can not be corrected through data assimilation. This may be confirmed through examining the forecast step of the reduced-order Kalman filter in full space. If neglecting model errors and associated covariance matrix $\mathbf{Q}_k$, the advancement in time of the covariance matrix for the full state may be written as

$$\mathbf{P}_k^f = \mathbf{A}\mathbf{P}_{k-1}^a\mathbf{A}^{\mathrm{T}} = \Phi\Lambda\Phi^{\dagger}\mathbf{P}_{k-1}^a\mathbf{A}^{\mathrm{T}} \tag{17}$$

The forecast step thus remains constrained in the subspace spanned by the DMD modes Φ. In addition, from the examination of the expression of the Kalman gain in (4) and that of the analysis step in (3), it appears that the assimilated state also remains confined to the DMD subspace.

In order to overcome this limitation, it is here proposed to first build a catalogue of DMDs that correspond to different values of the uncertain parameters (here the Reynolds number). These DMDs are then used to build an augmented subspace by taking the union of the DMDs. This augmented subspace should allow to better represent the true state compared to the use of any single DMD in the catalogue. For example, considering the case where we have access to two ensembles of snapshots $\{\mathbf{X}_1, \mathbf{X}_1'\}$ and $\{\mathbf{X}_2, \mathbf{X}_2'\}$, where the subscript refers to the value of the Reynolds number, the associated DMDs are first performed according to

$$\mathbf{A}_1 = \mathbf{X}_1' \mathbf{X}_1^\dagger, \quad \mathbf{A}_2 = \mathbf{X}_2' \mathbf{X}_2^\dagger \tag{18}$$

$$\mathbf{A}_1 \Phi_1 = \Phi_1 \Lambda_1, \quad \mathbf{A}_2 \Phi_2 = \Phi_2 \Lambda_2 \tag{19}$$

An augmented subspace Φ_{tot} may then be built according to

$$\Phi_{tot} = [\Phi_1 \ \Phi_2] \tag{20}$$

Reduced-order Kalman filtering may be performed based on this augmented subspace and from the following new definition of the reduced state

$$\tilde{\boldsymbol{x}} = \Phi_{tot}^\dagger \boldsymbol{x} \tag{21}$$

This new reduced state is advanced in time according to

$$\tilde{\boldsymbol{x}}_{k+1} = \Lambda_{tot} \tilde{\boldsymbol{x}}_k, \quad \Lambda_{tot} = \left[\begin{array}{cc} \Lambda_1 & \mathbf{0} \\ \mathbf{0} & \Lambda_2 \end{array} \right] \tag{22}$$

The above expressions are straightforwardly extended to the case where the catalogue is formed by more than two DMDs. The present Catalogue Kalman Filter (CKF), or Catalogue EnKF (CEnKF) when the EnKF is employed as is the case here, is assessed in next section.

6. Results for the Catalogue EnKF

The same test case and geometry will be used in this section to compare the results of the present CEnKF approach with the ones of section 4.2 that were obtained from single DMDs. The catalogue will consist of DMD modes obtained for the Reynolds numbers 98 and 102. 19 modes are generated for each Reynolds number and thus the catalogue is composed of 38 modes (two corresponding to the mean flow and 18 pairs of complex conjugate modes). The CEnKF will be performed with 38 ensemble members, as 19 will be generated from the DMD at $Re = 98$ and 19 others from the DMD at $Re = 102$. The true Reynolds number will first be set at 100 in order to allow comparisons with results of section 4.2. Then, the true state will correspond to $Re = 105$, which is outside the catalogue, in order to test the extrapolation capabilities of the proposed methodology in parameter space.

6.1 Truth at Reynolds 100

Considering a reference flow to reconstruct at $Re = 100$, as in section 4.2, reconstruction results obtained by the present CEnKF methodology are first illustrated in figures 9a-9c. It appears from the comparison between these figures and figure 6 that the use of the CEnKF has significantly decreased the errors in the estimated vorticity fields compared to the use of any of the DMDs in the catalogue. A "DMD update" is then performed, i.e. a new DMD is computed based on the last 400 assimilated states, and reduced-order Kalman filtering is then applied in the corresponding reduced-order space. This filtering is performed in the same conditions as before the update: 38 ensemble members are used and 37 modes are calculated (before the update there were 38 modes in which 2 corresponded to the mean flow and were not in a complex conjugate pair). The corresponding reconstruction results are illustrated in figures 9d-9f. This DMD update seems to provide some slight improvement compared to the baseline CEnKF results in figures 9a-9c.

The time evolution of the error E_k for these reconstruction results from the CEnKF is reported in dash-dotted lines in figure 10. For the sake of comparison, results of section 4.2 obtained when relying on a single DMD at $Re = 98$ or $Re = 102$ (dashed and dotted lines), or at the true Reynolds number $Re = 100$ (full line), are reproduced. The CEnKF approach has allowed to decrease by one order of magnitude the error E_k compared to the use of any single DMD in the catalogue, reaching an error level that is close to the case when relying on a DMD built at the true Reynolds number.
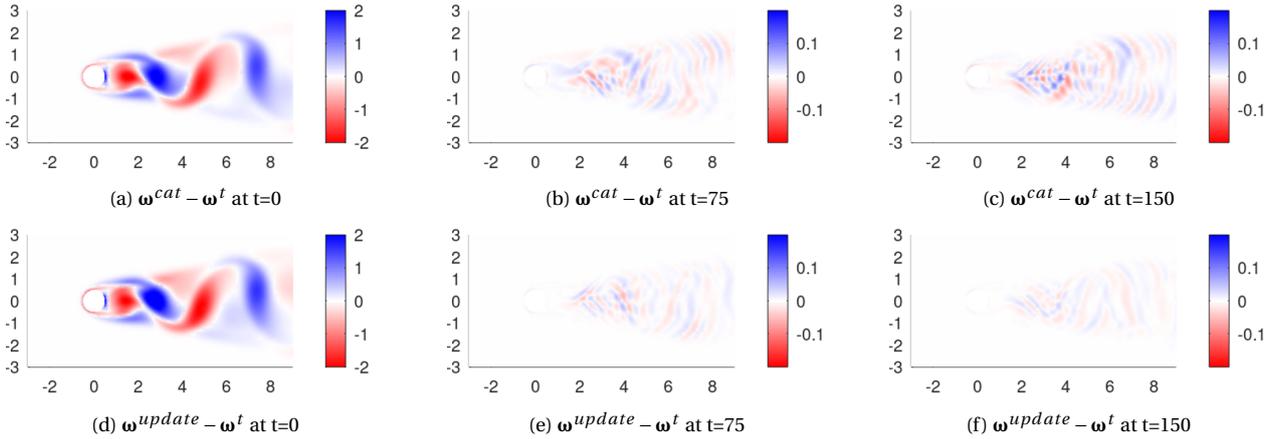
Figure 9: First and second rows: difference between estimated and true vorticity fields for the CEnKF (first row) and after an update of the DMD (second row). First to third column correspond to times $t = 0$, $t = 75$ and $t = 150$. The true flow corresponds to R$e = 100$.
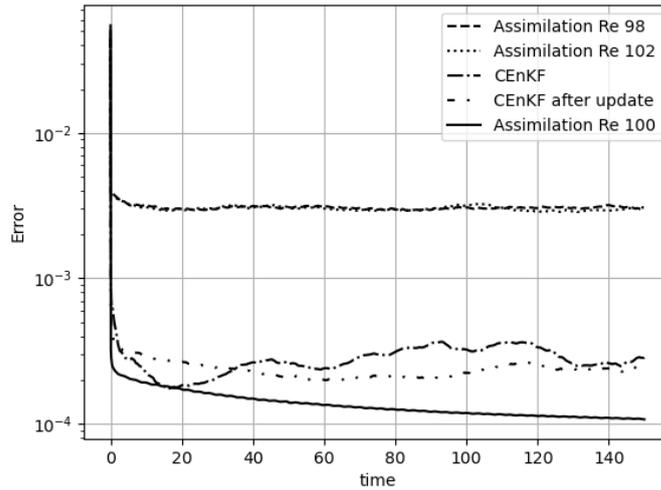
Figure 10: Temporal evolution of the error E_k for the reconstructed flow obtained with the CEnKF and after an update of the DMD (dash-dotted lines). E_k is also reported for reconstructed flows obtained from a single DMD at R$e = 98$, R$e = 102$ (dash and dotted lines), and at the true Reynolds number R$e = 100$ (full line).

To further analyse the performances of the CEnKF, the discrepancies between the DMD at the true Reynolds number and that performed on the reconstructed flow from the CEnKF are illustrated in figure 11 through full lines. The same discrepancies are reported for the DMDs that form the catalogue in dashed and dotted lines. Similarly as in section 4.2 where single DMDs were used, the CEnKF allows to correctly estimate the frequencies in the true flow, significantly decreasing errors compared to the DMDs in the catalogue, as illustrated by figure 11a. More interestingly, it appears from figure 11b that, contrary to the case where reduced-order Kalman filtering was performed based on single DMDs of the catalogue (see figure 8b), the assimilation procedure based on the CEnKF enables a significant improvement in the estimation of the spatial structure of the first true DMD modes. Errors appear higher compared to the first-guess DMD modes in the catalogue for some modes after the 11th one, but the latter do not correspond to dynamically important ones.

These results thus confirm the ability of the proposed CEnKF methodology in overcoming parameter uncertainties in the context of reduced-order Kalman filtering.

(a)   (b)
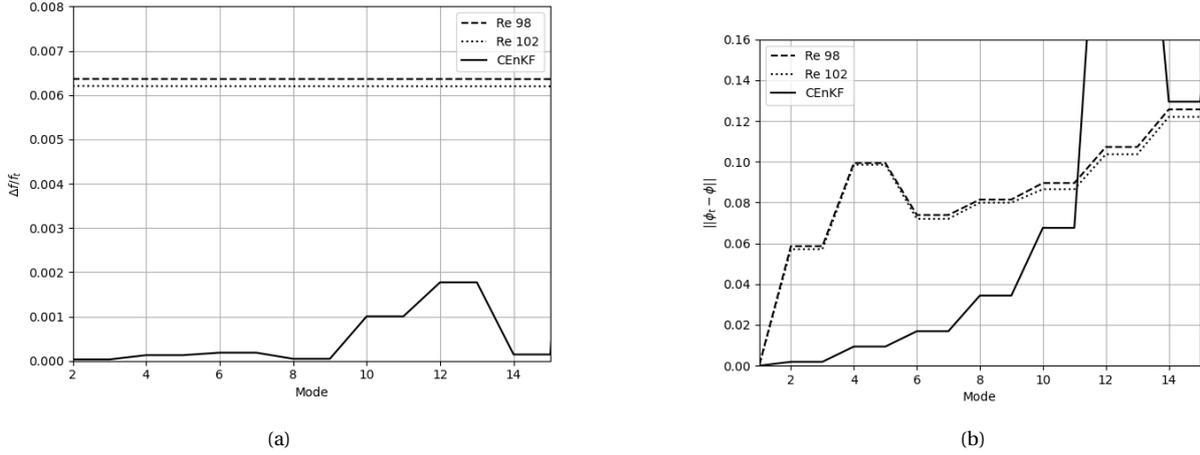
Figure 11: Error on (a) the frequency of DMD modes and (b) on the modes themselves for DMDs at Re = 98 and Re = 102 (dashed and dotted lines) and for the reconstructed flow that has been obtained with the CEnKF. The true flow corresponds to Re = 100.

6.2 Truth at Reynolds 105

Using the same catalogue of DMDs, we now investigate the reconstruction of a reference flow that corresponds to Re = 105, which is therefore outside of the interval of Reynolds numbers in the catalogue. First, reduced-order Kalman filtering is performed based on single DMDs in the catalogue. The CEnKF methodology is then applied. A supplementary assimilation is finally performed based on a updated DMD that is obtained from the CEnKF results.

The corresponding results are illustrated in figure 12. Errors in the estimated vorticity fields when reduced-order Kalman filtering is performed based on a single DMD at Re = 98 or Re = 102 appear large. Large-scale structures may even be identified in the error fields. Not surprisingly, errors are larger when relying on the DMD at Re = 98 compared to the use of that at Re = 102 in this case. The CEnKF approach appears to significantly improve the reconstruction results, while an update of the DMD from the so-obtained assimilated snapshots also provides satisfactory (but not necessarily better) results.

These findings are confirmed by figure 13, where is reported the error E$_k$ for these cases. Similarly as in section 6.1, the CEnKF methodology reaches an error level that is roughly one order of magnitude lower compared to the use of any single DMD in the catalogue.

Finally, the discrepancies between the DMD of the true flow and that obtained from the reconstructed flow by the CEnKF are illustrated in figure 14, comparing with the initial discrepancies for the DMDs in the catalogue. Again, in addition to the correct identification of the frequencies of the true flow, the CEnKF methodology has allowed to significantly enhance the estimation the spatial structure of the modes, putting aside modes beyond the 13th one which are not dynamically significant. The improvement in the estimation of the 4th mode is specifically illustrated in figure 15.

All these findings confirm the ability of the CEnKF in providing satisfactory reconstruction results, even when the flow to reconstruct corresponds to a parameter value that is outside from the interval in the catalogue.

7. Conclusion

In this paper, we have investigated Kalman filtering based on reduced-order models, here relying on DMD, to perform data assimilation and flow reconstruction at low computational cost. In order to tackle parameter uncertainties when forming the reduced-order model, it has been proposed to form a catalogue of DMDs calculated for different values of the uncertain parameters and build an augmented subspace from the union of the DMDs.

The present Catalogue (Ensemble) Kalman Filter (C(En)KF) has been assessed for flow reconstruction of the flow past a cylinder from pointwise measurements, considering the Reynolds number of the flow as badly-known. It has been confirmed that this methodology allows to provide significantly better reconstruction results compared to the use of any single DMD in the catalogue, optimally combining the different DMD bases to estimated both the frequencies and the spatial structures of the flow. The robustness of the CEnKF has been evaluated considering a true flow whose Reynolds number is in the interval of the catalogue or not.
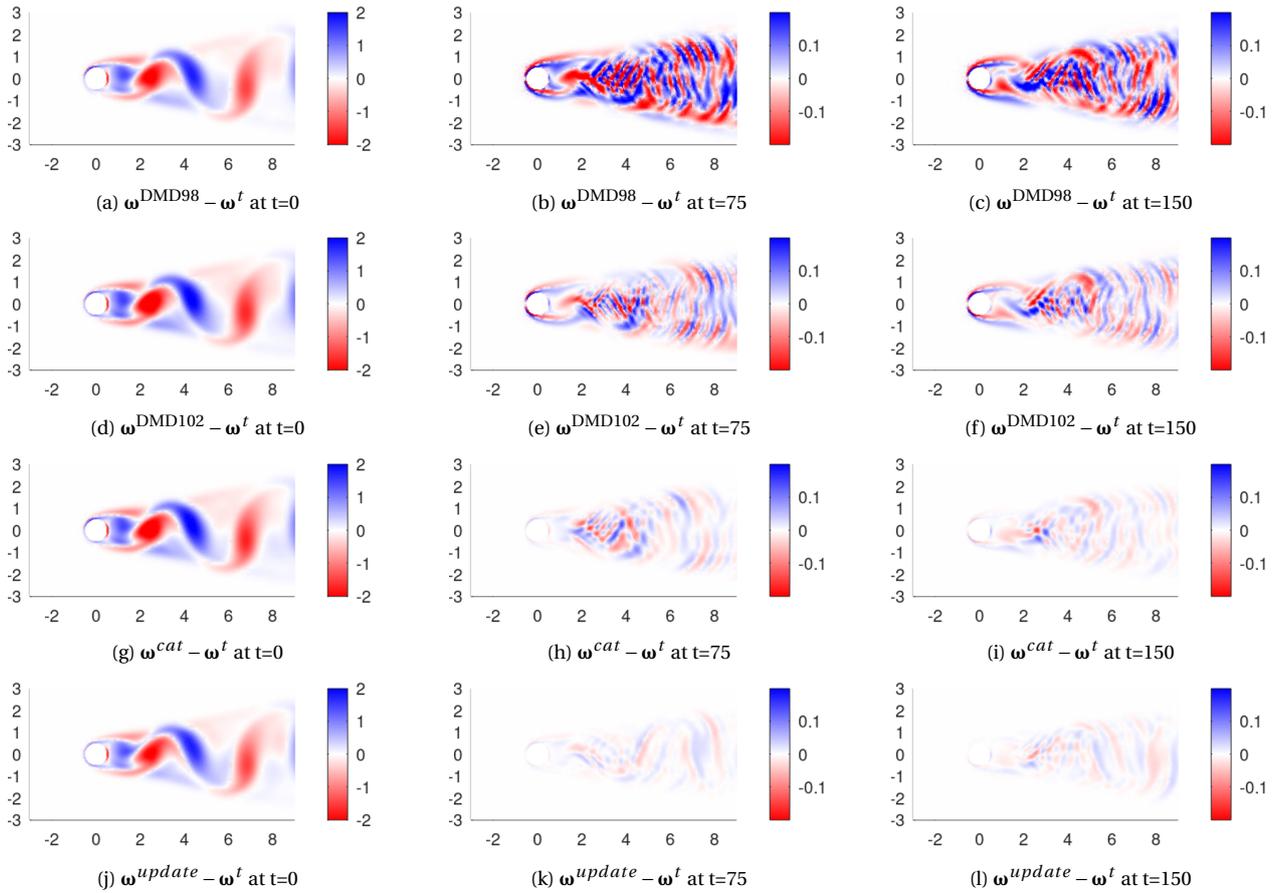
Figure 12: First to fourth rows: difference between estimated and true vorticity fields when the reduced-order EnKF is based on a DMD at Re = 98 (first row) or Re = 102 (second row), for the CEnKF (third row) and after an update of the DMD (fourth row). First to third column correspond to times t = 0, t = 75 and t = 150. The true flow corresponds to Re = 105.
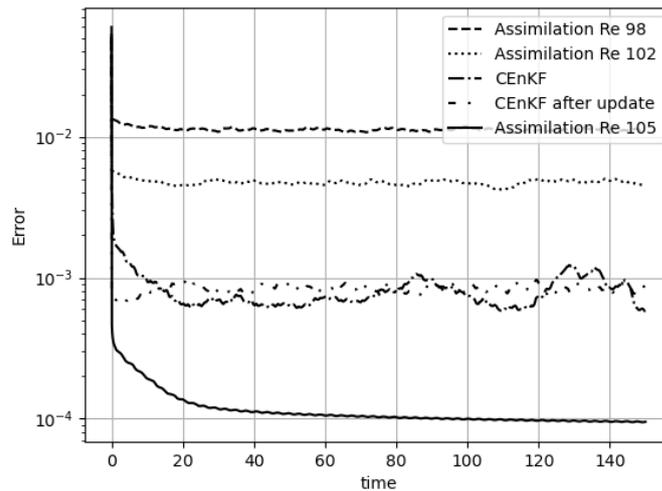
Figure 13: Temporal evolution of the error E$_k$ for the reconstructed flow obtained with the CEnKF and after an update of the DMD (dash-dotted lines). E$_k$ is also reported for reconstructed flows obtained from a single DMD at Re = 98, Re = 102 (dash and dotted lines), and at the true Reynolds number Re = 105 (full line).

(a)   (b)
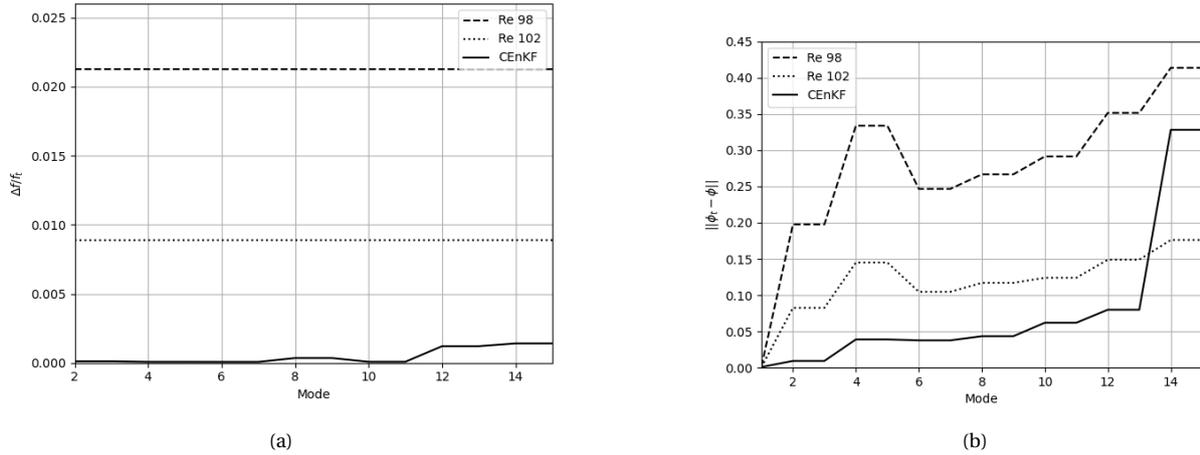
Figure 14: Error on (a) the frequency of DMD modes and (b) on the modes themselves for DMDs at Re = 98 and Re = 102 (dashed and dotted lines) and for the reconstructed flow that has been obtained with the CEnKF. The true flow corresponds to Re = 105.

(a)   (b)   (c)
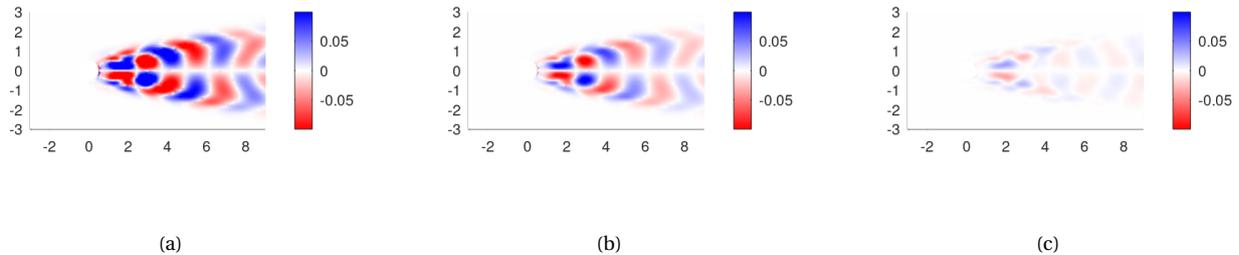
Figure 15: Vorticity field for the real part of the difference between the true 4th DMD mode at Re = 105 and the one for the DMDs at (a) Re = 98, (b) Re = 102 and (c) from the reconstructed flow with the CEnKF.

Extensions of the present methodology include the re-estimation of the parameter of the true flow from the reconstructed flow by the CEnKF. The present methodology could also be further validated considering more complex flows and with more uncertain parameters. Finally, it would be interesting to apply the CEnKF to actual experimental data.

References

[1] K. K. Chen, J. H. Tu, and C. W. Rowley. Variants of Dynamic Mode Decomposition: Boundary Condition, Koopman, and Fourier Analyses. *Journal of Nonlinear Science*, 22:887–915, 2012.

[2] Geir Evensen. *Data assimilation: the ensemble Kalman filter*. Springer Science & Business Media, 2009.

[3] Mojtaba F. Fathi, Ahmadreza Baghaie, Ali Bakhshinejad, Raphael H. Sacho, and Roshan M. D'Souza. Time-resolved denoising using model order reduction, dynamic mode decomposition, and Kalman filter and smoother. *Journal of Computational Dynamics*, 7(2):469–487, 2020.

[4] T. Hayase. Numerical simulation of real-world flows. *Fluid Dynamics Research*, 47:051201, 2015.

[5] F. Hecht. New development in FreeFem++. *Journal of Numerical Mathematics*, 20:251–265, 2012.

[6] Giacomo V Iungo, Christian Santoni-Ortiz, Mahdi Abkar, Fernando Porté-Agel, Mario A Rotea, and Stefano Leonardi. Data-driven reduced order model for prediction of wind turbine wakes. In *Journal of Physics: Conference Series*, volume 625, page 012009. IOP Publishing, 2015.

[7] Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. 1960.

A DMD APPROACH FOR DATA ASSIMILATION

[8] Jiaqing Kou and Weiwei Zhang. An improved criterion to select dominant modes from dynamic mode decomposition. *European Journal of Mechanics - B/Fluids*, 62, 12 2016.

[9] Patrick N. Raanes, Colin Grudzien, and 14tondeu. nansencenter/dapper: Version 0.8, December 2018.

[10] Peter J. Schmid. Dynamic mode decomposition of numerical and experimental data. *Journal of Fluid Mechanics*, 656:5 – 28, 2010.

[11] Gilles Tissot, Laurent Cordier, Nicolas Benard, and Bernd R Noack. Model reduction using dynamic mode decomposition. *Comptes Rendus Mécanique*, 342(6-7):410–416, 2014.

[12] Alexandros Tsolovikos, Saikishan Suryanarayanan, Efstathios Bakolas, and David B. Goldstein. Multiple model dynamic mode decomposition for flowfield and model parameter estimation. *AIAA SCITECH 2022 Forum*.

[13] Jonathan H. Tu, Clarence W. Rowley, Dirk M. Luchtenburg, Steven L. Brunton, and J. Nathan Kutz. On dynamic mode decomposition: Theory and applications. *Journal of Computational Dynamics*, 1(2):391–421, 2014.