Dense velocity, pressure and Eulerian acceleration fields from single-instant scattered velocities through Navier–Stokes-based data assimilation

Vincent Mons, Olivier Marquet, Benjamin Leclaire, Philippe Cornic, Frédéric Champagnat

Dense velocity, pressure and Eulerian acceleration fields from single-instant scattered velocities through Navier-Stokes-based data assimilation

V. Mons[1], O. Marquet[1], B. Leclaire[1], P. Cornic[2] and F. Champagnat[2]

[1]DAAA, ONERA, Université Paris Saclay, F-92190 Meudon, France
[2]DTIS, ONERA, Université Paris Saclay, F-91123 Palaiseau, France

E-mail: `vincent.mons@onera.fr`

July 2022

Abstract. In this study, a reconstruction procedure to infer full 3D instantaneous velocity and pressure fields from sparse velocity measurements is proposed, here focusing on the case of scattered data as provided by Particle Tracking Velocimetry (PTV). A key characteristic of the present approach is that it only relies on single-instant velocity measurements, and does not require any time-resolved or acceleration information. It is based on a strong enforcement of the Navier-Stokes equations where the partial time derivative of the velocity field, namely Eulerian acceleration, is considered as a control vector to minimize the discrepancies between the single-instant measurements and the reconstructed flow. Eulerian acceleration is thus a byproduct of the present methodology in addition to the identification of the full velocity and pressure fields. The reconstruction performances of the proposed Navier-Stokes-based data-assimilation approach for single-instant velocity measurements (NS-DA-SIM) are assessed using a numerical dataset for the 3D flow past a cylinder. Comparisons with existing data assimilation methodologies allow to further illustrate the merits of the present approach. The latter is finally applied to the instantaneous reconstruction of an experimental air jet flow from volumetric PTV data, confirming its robustness and high efficacy.

1. Introduction

Given the potential high complexity of flows in engineering applications, especially in turbulent regimes, a large majority of the current well-established measurement techniques in fluid mechanics are still characterized by the necessity of a trade-off between spatial (and/or temporal) resolution, sampling, and accuracy. Emblematic examples are volumetric particle-based velocimetry techniques, which have been the subject of important research efforts in the past fifteen years. While tomographic Particle Image Velocimetry (PIV) [1] can provide 3D velocity fields on Cartesian grids, allowing a straightforward evaluation of velocity gradients and flow statistics, it is now

well-known that these fields oftentimes suffer a strong spatial filtering. This is due to the seeding density that is lower in tomographic PIV than in planar PIV. On the other hand, 3D Particle Tracking Velocimetry (PTV) can provide much more accurate velocity estimations, with maximum reachable seeding densities that are comparable to that of tomographic PIV (see, e.g., [2, 3]). This comes at the cost of obtaining a scattered vector field that coincides with the particles' positions; besides, the maximum possible seeding density may still be too low for satisfactorily characterizing high-Reynolds-number turbulent flows. It therefore appears difficult to obtain accurate, dense and regularly sampled (in the spatial sense) instantaneous velocity fields from volumetric optical velocimetry alone, not to mention the inference of other flow quantities such as pressure [4].

In the present study, we are interested in obtaining a dense and full description of the instantaneous flow observed here with 3D PTV, namely accurate but sparse velocity measurements. While this aim may be achieved through many approaches, including the use of advanced interpolation schemes or that of model-reduction techniques [5, 6], or even recent machine learning methodologies [7, 8], we here investigate the use of physical constraints/conservation laws to perform the reconstruction of the full flow. Such an approach may thus be referred to as data assimilation [9]. Furthermore, we aim to perform data assimilation relying on *single-instant velocity measurements only*, i.e. without requiring the acquisition of time-resolved data. The proposed method is thus suited to a large range of experiments, as it might be applied on PTV data obtained with high as well as low frame rate laser and camera systems. Besides, although we here demonstrate it in the context of volumetric PTV measurements, it could be as well directly employed with other multi-point measurements.

In this context and in the case of incompressible flows, as will be considered here, a first physical constraint to take into account may be the divergence-free character of the velocity field. This flow property was exploited in many previous studies. Among others, it was included in interpolation schemes (based on radial basis functions [10] or Kriging [11]) so that the outcome is automatically divergence-free. This condition was also used in variational approaches, where the flow reconstruction problem is formulated as the minimization of a cost function that evaluates the discrepancies between measurements and the reconstructed velocity field, the divergence-free condition being considered as an equality constraint (see, e.g., [12]). Beyond velocity, methodologies to infer pressure from single-instant PIV fields were proposed in [13, 14], where a Poisson problem for the pressure field is solved. This problem is forced by the Lagrangian acceleration (velocity material derivative) [4], which is determined in these studies through the advancement over a short time interval of a vorticity field that is directly evaluated from the measured velocity field, relying on the vorticity transport equation.

In order to further enhance the fidelity of the regularized flow and favor super-resolution, one may consider the imposition of more sophisticated conservation laws, namely the full (and coupled) unsteady Navier-Stokes equations. Based on the argument that such a regularization could be implemented most straightforwardly in conjunction

with time-resolved measurements, the use of the Navier-Stokes equations was mainly investigated in this context. Among other approaches, they were employed in [15] to fill gaps in PIV measurements, the measurements serving as unsteady boundary conditions in this procedure. Various data assimilation techniques were also developed to inject PTV measurements into Navier-Stokes simulations through a feedback term in the momentum equation, ranging from simple state-observer/nudging schemes to Kalman filter techniques [16]. Furthermore, variational methodologies were extended to the use of the Navier-Stokes equations, either in their standard form or considering variants of the vorticity equation. This was notably performed for PTV in the approaches proposed in [17, 18, 19], which rely on both velocity *and* Lagrangian acceleration data, therefore requiring time-resolved measurements. Furthermore, while the latter techniques may provide satisfactory estimations of velocity and pressure fields, they involve cost functions that include multiple data and penalization terms, whose respective weights may not be straightforward to prescribe. This may be at least partially alleviated through a strong enforcement of the unsteady Navier-Stokes equations, generally relying on measurement data at multiple instants within a given time horizon, as performed in [20, 21, 22, 12, 23] for PIV measurements and in [24] in the case of PTV. This approach possibly corresponds to the most elaborate and rigorous variational data assimilation methodology for unsteady measurements, the consideration of single-instant measurements only in this framework having been investigated in [25], and will be here revisited for comparisons with the proposed methodology in this study. However, unsteady variational data assimilation is challenging to implement, as it requires solving the adjoint of the unsteady linearized Navier-Stokes equations, which needs the storage or recomputation of the full history of the estimated flow over the considered time window. Besides, maintaining the computational cost of such an approach tractable would impose to consider a limited time-horizon, making it then necessary to consider high frame rate measurements, which are not always easily accessible in practice (hardware available, experimental conditions compatible with a lower signal-to-noise ratio, sufficiently low flow frequencies, for example).

We here present a novel variational data assimilation methodology for instantaneous 3D flow reconstruction from single-instant scattered velocity measurements, as in PTV, that does not require additional time-resolved information or Lagrangian acceleration data. To avoid the difficulties and limitations in considering fully unsteady variational data assimilation, we propose to use the time-independent (instead of unsteady) Navier-Stokes equations as a numerical model. We will show that this somehow surprising choice for the numerical operator does not prevent from reconstructing a flow snapshot since the Eulerian acceleration, i.e. the partial time derivative of the velocity field, is treated in the method as an unknown forcing in the momentum equation. It is then identified to minimize a cost function that quantifies the discrepancies between the single-instant velocity measurements and the reconstructed flow. While the proposed numerical model is similar to that used in previous works on mean flow reconstruction [26, 27], the output solution and the input forcing have different physical meanings in both methods. As

stated above, they here represent the instantaneous flow and the Eulerian acceleration, respectively, rather than the time-averaged flow and the Reynolds stresses in [26, 27]. Interestingly, we will show that, as Eulerian acceleration is a divergence-free quantity, the proposed approach allows reconstructing the pressure field, without suffering from the deficiency identified in [26, 27] for mean flow reconstruction. Compared to classical unsteady variational approaches that require forward/backward time integration, a benefit of the present method is that it only requires a steady adjoint model. We will demonstrate, through both numerical and experimental test cases, that the proposed Navier-Stokes-based Data-Assimilation approach for Single-Instant velocity Measurements (NS-DA-SIM) enables faithfully reconstructing the full instantaneous velocity along with providing as well the pressure and Eulerian acceleration fields, which were not even accessible a priori at the measurement locations. Furthermore, the scarcity of the velocity data is here tackled in the optimization procedure through a H^1-like regularization of the gradient of the cost function with respect to the control vector (here the Eulerian acceleration) [28, 29, 30], instead of adding terms in the cost function that would penalize its spatial gradients and introduce tunable parameters. The smoothing scale that is involved in this regularization technique may be here straightforwardly related to the average distance between neighboring measurements. As such, and in conjunction with the strong enforcement of the Navier-Stokes equations, the optimization problem in the proposed approach is free of any adjustable parameter, thus fostering its application to a variety of configurations.

The paper is organized as follows. The proposed NS-DA-SIM methodology is first described in §2. The generation of a first guess for the optimization procedure is detailed in §3, along with the H^1-like regularization of the gradient and numerical methods. The approach is then assessed, first in §4 by relying on synthetic observations that are extracted from the direct numerical simulation of the 3D flow past a cylinder at $Re = 300$, and then in §5 by considering PTV measurements of an air jet flow at $Re = 4600$ [3]. Finally, concluding remarks and perspectives are drawn in §6.

2. Navier-Stokes-based data assimilation of single-instant velocity measurements (NS-DA-SIM)

2.1. Formulation of the NS-DA-SIM approach

We here investigate the reconstruction of incompressible Newtonian flows. All quantities in this section and in the following are supposed already non-dimensionalized based on reference velocity U, length scale D, and the fluid density ρ, which will be specified for the considered flows in §4-§5. The velocity measurements which we consider in the following correspond to a single-instant $t = t_m$ and are denoted by $\boldsymbol{m}(t_m)$. We will here particularly investigate the case of 3D scattered data as may be provided, e.g., through 3D PTV. In this context, $\boldsymbol{m}(t_m)$ corresponds to a vector of size $3N_{\mathrm{PTV}}$, with N_{PTV} the

number of 3D velocity vector data, and may be detailed as

$$\boldsymbol{m}(t_m) = (\boldsymbol{m}_1(t_m)^{\mathrm{T}}, \boldsymbol{m}_2(t_m)^{\mathrm{T}}, \cdots, \boldsymbol{m}_{N_{\mathrm{PTV}}}(t_m)^{\mathrm{T}})^{\mathrm{T}}, \tag{1}$$

with $\boldsymbol{m}_i(t_m)$ the i-th 3D velocity measurement at location $\boldsymbol{x}_i$ (particle location in the case of PTV). The density of the measurements will be here characterized by the mean distance between neighboring measurement locations, $\overline{d}$, so that dense measurements correspond to low $\overline{d}$ and vice versa. In the case of PTV measurements, $\overline{d}$ thus identifies with the mean inter-particle distance.

In the following, we aim to infer the full instantaneous flow field that is associated to the measurements $\boldsymbol{m}(t_m)$. The flow of interest is solution of the Navier-Stokes equations, which may be written in the following non-dimensional form

$$\frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} - Re^{-1}\Delta\boldsymbol{u} + \nabla p = \boldsymbol{0}, \quad \nabla\cdot\boldsymbol{u} = 0, \tag{2}$$

where $\boldsymbol{u}$ and p are the velocity and pressure fields, respectively, $Re = UD/\nu$ is the Reynolds number and ν the kinematic viscosity. We here thus aim to reconstruct $\boldsymbol{u}(t_m)$ and $p(t_m)$ from $\boldsymbol{m}(t_m)$. Even if only single-instant measurements are supposed to be available, one could consider directly relying on the unsteady Navier-Stokes equations (2) to perform variational data assimilation (see Appendix A.2), as investigated in the following (see §4.3). Instead, one may first write the momentum equation evaluated at the measurement time t_m according to

$$\left((\boldsymbol{u}\cdot\nabla)\boldsymbol{u} - Re^{-1}\Delta\boldsymbol{u} + \nabla p\right)|_{t_m} = -\frac{\partial \boldsymbol{u}}{\partial t}\Big|_{t_m} = \boldsymbol{f}(t_m). \tag{3}$$

The key proposal that underlies the present methodology is to adopt the viewpoint that (3) forms steady equations. In this perspective, the opposite of the Eulerian acceleration $\frac{\partial \boldsymbol{u}}{\partial t}|_{t_m}$ is treated as a forcing $\boldsymbol{f}(t_m)$ to these equations. Dropping the notation that all quantities correspond to the measurement time t_m for the sake of simplicity in the following, (3) and the mass conservation equation may be rewritten in the following compact form

$$\boldsymbol{\mathcal{N}}(\boldsymbol{q}) = \mathbf{P}\boldsymbol{f}, \tag{4}$$

where $\boldsymbol{q} = (\boldsymbol{u}^{\mathrm{T}}, p)^{\mathrm{T}}$ is the full flow field, or state-space vector, while the nonlinear operator $\boldsymbol{\mathcal{N}}$ and linear operator $\mathbf{P}$ are defined through

$$\boldsymbol{\mathcal{N}}(\boldsymbol{q}) = \begin{pmatrix} (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} - Re^{-1}\Delta\boldsymbol{u} + \nabla p \\ -\nabla\cdot\boldsymbol{u} \end{pmatrix}, \quad \mathbf{P} = \begin{pmatrix} \mathbb{I} \\ 0 \end{pmatrix}, \tag{5}$$

where $\mathbb{I}$ is the identity matrix. The action of $\mathbf{P}$ thus makes the forcing $\boldsymbol{f}$ act on the momentum equation only. A variational data assimilation procedure to reconstruct the full flow from the single-instant velocity measurements $\boldsymbol{m}$ based on the imposition of the steady Navier-Stokes equations (4) may be designed as follows. A fortiori in the case of single-instant measurements, the Eulerian acceleration, and so $\boldsymbol{f}$, is unknown. Accordingly, it is here proposed to formulate the data assimilation problem as identifying the forcing $\boldsymbol{f}$ that allows to minimize the discrepancies between the measurements $\boldsymbol{m}$

and the reconstructed velocity field $\boldsymbol{u}$. The corresponding optimization problem, which forms the core of the present NS-DA-SIM approach, may be written as

$$\min_{\boldsymbol{f}} \left\{ J = \frac{1}{2} \|\boldsymbol{m} - \mathbf{H}\boldsymbol{u}\|^2_{\mathcal{M}} \right\}, \quad \boldsymbol{\mathcal{N}}(\boldsymbol{q}) = \mathbf{P}\boldsymbol{f}, \tag{6}$$

where $\mathbf{H}$ is a linear operator that maps the model space to the measurement one. In the present case, if we assume the data $(\boldsymbol{m}_i)_{i\in\{1,2,\cdots,N_{\mathrm{PTV}}\}}$ in (1) to correspond to strictly pointwise measurements of the velocity field of interest, the application of $\mathbf{H}$ on the velocity field $\boldsymbol{u}$ simply consists in evaluating the latter at the measurement locations according to

$$\mathbf{H}\boldsymbol{u} = (\boldsymbol{u}(\boldsymbol{x}_1)^{\mathrm{T}}, \boldsymbol{u}(\boldsymbol{x}_2)^{\mathrm{T}}, \cdots, \boldsymbol{u}(\boldsymbol{x}_{N_{\mathrm{PTV}}})^{\mathrm{T}})^{\mathrm{T}}. \tag{7}$$

The norm used in the expression of the cost function J in (6) corresponds to the Euclidean norm on $\mathbb{R}^{3N_{\mathrm{PTV}}}$, i.e. $\|\boldsymbol{m} - \mathbf{H}\boldsymbol{u}\|^2_{\mathcal{M}} = \sum_{i=1}^{N_{\mathrm{PTV}}} (\boldsymbol{m}_i - \boldsymbol{u}(\boldsymbol{x}_i)) \cdot (\boldsymbol{m}_i - \boldsymbol{u}(\boldsymbol{x}_i))$.

Before detailing how the equality-constrained minimization problem (6) is solved in §2.2, two further comments may be made about the expression of the cost function J. Firstly, measurement errors could be taken into account by replacing the standard Euclidean norm with the definition $\|\boldsymbol{a}\|^2_{\mathcal{M}} = \boldsymbol{a}^{\mathrm{T}}\mathbf{C}\boldsymbol{a}$, where $\boldsymbol{a} \in \mathbb{R}^{3N_{\mathrm{PTV}}}$ and $\mathbf{C}$ is a symmetric $3N_{\mathrm{PTV}} \times 3N_{\mathrm{PTV}}$ matrix which should correspond to the inverse of the measurement error covariance matrix, assuming Gaussian statistics [9]. In the case of uncorrelated, isotropic and homogeneous errors, $\mathbf{C}$ should reduce to $\mathbf{C} = c\mathbb{I}$ with c a positive real number. This would amount to multiply by c the present expression of J, which should thus provide the same optimization results as with the present formulation. However, in the case where finer information about measurement errors are available, such information could thus be straightforwardly included in the present approach. Secondly, as $\boldsymbol{m}$ here corresponds to scattered velocity measurements, one might be surprised that the expression of J does not include any regularization/penalty term on the forcing $\boldsymbol{f}$. This point is in fact accounted for in a different manner here and will be addressed in §3.2.

2.2. Adjoint-based optimization procedure

The equality-constrained minimization problem (6) in the present NS-DA-SIM approach is solved following the adjoint approach [31, 9], which relies on the transformation of (6) into an unconstrained optimization problem through the introduction of the following Lagrangian

$$\mathcal{L} = J - \left\langle \boldsymbol{q}^\dagger, \boldsymbol{\mathcal{N}}(\boldsymbol{q}) - \mathbf{P}\boldsymbol{f} \right\rangle_{\mathcal{Q}}, \tag{8}$$

where $\boldsymbol{q}^\dagger = (\boldsymbol{u}^{\dagger\,\mathrm{T}}, p^\dagger)^{\mathrm{T}}$ is referred to as the adjoint state, while the involved scalar product is defined as $\langle \boldsymbol{a}, \boldsymbol{b} \rangle_{\mathcal{Q}} = \int_\Omega \boldsymbol{a} \cdot \boldsymbol{b}\, \mathrm{d}\Omega$, where Ω is the flow domain and $\boldsymbol{a}$ and $\boldsymbol{b}$ are two vectorial functions of space. In this framework, the solution of the minimization problem should cancel the partial variations of $\mathcal{L}$ with respect to the direct state $\boldsymbol{q}$, the adjoint state $\boldsymbol{q}^\dagger$ and the forcing $\boldsymbol{f}$. Cancelling variations with respect to $\boldsymbol{q}^\dagger$ gives back

the steady Navier-Stokes equations (4). Cancelling variations with respect to $\boldsymbol{q}$ provides the following linear problem for the adjoint state

$$\mathbf{N}^\dagger \boldsymbol{q}^\dagger = \mathbf{P}\mathbf{H}^\dagger(\mathbf{H}\boldsymbol{u} - \boldsymbol{m}), \tag{9}$$

where the adjoint Navier-Stokes operator $\mathbf{N}^\dagger$ is defined as

$$\mathbf{N}^\dagger = \begin{pmatrix} \circ \cdot (\nabla \boldsymbol{u})^{\mathrm{T}} - (\boldsymbol{u} \cdot \nabla) \circ - Re^{-1}\Delta \circ & -\nabla \circ \\ \nabla \cdot \circ & 0 \end{pmatrix}, \tag{10}$$

while the right-hand-side involves the adjoint observation operator $\mathbf{H}^\dagger$ and may be detailed as

$$\mathbf{H}^\dagger(\mathbf{H}\boldsymbol{u} - \boldsymbol{m}) = \sum_{i=1}^{N_{\mathrm{PTV}}} \delta(\boldsymbol{x} - \boldsymbol{x}_i)(\boldsymbol{u}(\boldsymbol{x}_i) - \boldsymbol{m}_i). \tag{11}$$

The determination of the above expressions relies on the following definition of adjoint operator: considering a linear operator $\mathbf{F}$ from vector space $\mathcal{A}$ to vector space $\mathcal{B}$, its adjoint operator $\mathbf{F}^\dagger$ verifies $\langle \boldsymbol{b}, \mathbf{F}\boldsymbol{a} \rangle_{\mathcal{B}} = \left\langle \mathbf{F}^\dagger \boldsymbol{b}, \boldsymbol{a} \right\rangle_{\mathcal{A}} \ \forall\, \boldsymbol{a} \in \mathcal{A}, \forall\, \boldsymbol{b} \in \mathcal{B}$. From solving the adjoint problem (9), one may evaluate the partial derivative of $\mathcal{L}$ with respect to $\boldsymbol{f}$, or equivalently the total derivative of the cost function J with respect to $\boldsymbol{f}$ according to

$$\frac{dJ}{d\boldsymbol{f}} = \boldsymbol{u}^\dagger. \tag{12}$$

More details about the above derivations may be found in [26, 32]. From these expressions, one may employ an iterative gradient-based descent method to solve (6) which may be summarized as follows:

(i) Start with a first guess for the forcing $\boldsymbol{f}$. The generation of the latter is detailed in §3.1.

(ii) Solve the steady Navier-Stokes equations (4) (see §2.3 and §3.3). This provides a state $\boldsymbol{q}$.

(iii) Solve the adjoint problem (9). This provides an adjoint state $\boldsymbol{q}^\dagger$.

(iv) Compute the gradient $\frac{dJ}{d\boldsymbol{f}}$ according to (12). The latter is regularized following the procedure in §3.2.

(v) Update $\boldsymbol{f}$ according to

$$\boldsymbol{f} \leftarrow \boldsymbol{f} + s\boldsymbol{d}, \tag{13}$$

where the descent direction $\boldsymbol{d}$ is computed from the gradient from previous step based on the low-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) method [33], while the step size s is determined through the backtracking-Armijo line search algorithm [34].

(vi) Return to step (ii) until convergence.

The output of the above procedure is thus full instantaneous velocity and pressure fields, along with Eulerian acceleration through the optimal forcing $\boldsymbol{f}$. The corresponding flow will be referred to as the reconstructed flow in the following.

2.3. On boundary conditions

While the computational domain and boundary conditions to solve (4) and (9) for the considered flow configurations will be detailed in §4-§5, some general comments about the latter are here provided. First concerning the computational domain, it may be chosen indifferently to coincide or not with the measurement domain, which will be confirmed in §4.4. Concerning boundary conditions, it may be worth noting that the steady Navier-Stokes equations (4) are here solved in a fully coupled way, so that only boundary conditions for the velocity field $\boldsymbol{u}$ are required and imposed. Boundaries may be categorized in two groups: boundaries that correspond to easily identifiable boundary conditions (e.g. no-slip conditions at solid surfaces or uniform inflow conditions), and boundaries that do not. This latter situation may be encountered when, for example, an inlet boundary intersects a fluctuating part of the flow. In this case, the Dirichlet boundary condition $\boldsymbol{u} = \boldsymbol{u}^f$ is imposed, where $\boldsymbol{u}^f$ refers to the first-guess velocity field that is obtained in §3.1. The relevance of this choice will be confirmed by the quality of the reconstruction results in §4-§5. Boundary conditions for the adjoint problem (9) are derived from those imposed for the Navier-Stokes equations and from the integrations by parts that are performed to get $\mathbf{N}^\dagger$ based on the definition of adjoint operator (see §2.2). More details may be found in [26].

2.4. On pressure reconstruction

In this section, we discuss the ability of the present NS-DA-SIM methodology in correctly reconstructing pressure from velocity measurements alone, in contrast with other approaches. We first notice the similarities between the present methodology and the procedure for mean flow reconstruction proposed in [26]. The same steady Navier-Stokes equations (4) were solved in this work, and the data assimilation procedure also consisted in optimizing a forcing in these equations. The key difference between the present approach and [26] lies in the interpretation of the flow variables $\boldsymbol{q}$ and the forcing $\boldsymbol{f}$. In [26], $\boldsymbol{q}$ corresponded to mean-flow fields and $\boldsymbol{f}$ accounted for the opposite of the divergence of the Reynolds-stress tensor $-\nabla\cdot\mathbf{R}$. In the present case, $\boldsymbol{q}$ thus refers to instantaneous flow variables, while $\boldsymbol{f}$ should identify with $-\frac{\partial \boldsymbol{u}}{\partial t}$.

The difference in nature between $\nabla\cdot\mathbf{R}$ and $\frac{\partial \boldsymbol{u}}{\partial t}$ is determinant in the ability of data assimilation in correctly reconstructing the pressure field. Indeed, a variational data assimilation procedure where the incompressible Navier-Stokes (or Stokes, see Appendix A.1) equations are imposed can only reconstruct divergence-free forcings in the momentum equation if only velocity measurements are available. This may be inferred from (9)-(12), which show that $\nabla\cdot\frac{dJ}{d\boldsymbol{f}} = \nabla\cdot\boldsymbol{u}^\dagger = 0$, which entails that the optimal forcing obtained through the L-BFGS approach (see step (v) in §2.2) is divergence-free [35]. $\nabla\cdot\mathbf{R}$ is not a divergence-free vector, accordingly its Helmholtz decomposition may be written as $\nabla\cdot\mathbf{R} = \boldsymbol{f}^s + \nabla\phi$, where $\boldsymbol{f}^s$ is the divergence-free component and ϕ a (non-zero) potential. In the case of [26], only the component $\boldsymbol{f}^s$ could therefore be reconstructed. As a consequence, the data assimilation procedure could only give access

to a pseudo-pressure field $\pi = p + \phi$ with the impossibility of distinguishing between the true pressure p and the potential ϕ. In contrast, in the present case, the forcing $\boldsymbol{f}$ has to account for $-\frac{\partial \boldsymbol{u}}{\partial t}$, which is a divergence-free quantity. This implies that the latter may be reconstructed in its entirety and that the present NS-DA-SIM approach enables unambiguous and accurate estimation of the pressure field from velocity measurements alone, as confirmed in the following.

3. Implementation details

3.1. First-guess flow

The generation of a first guess for the forcing $\boldsymbol{f}$ (see step (i) in §2.2) is performed through the three following steps which may be summarized as: computing a first guess for the velocity field, then for the pressure field, and finally for the forcing $\boldsymbol{f}$.

Firstly, a first guess for the velocity field $\boldsymbol{u}$ is computed from the measurements $\boldsymbol{m}$ through a simple variational approach where we solely impose the divergence-free character of $\boldsymbol{u}$, while still enabling a prescribed degree of smoothness through penalization of the velocity gradients, similarly as considered in, e.g., [12]. $\boldsymbol{u}$ is thus sought as the solution of the following minimization problem

$$\min_{\boldsymbol{u}} \left\{ \frac{1}{2} \|\boldsymbol{m} - \mathbf{H}\boldsymbol{u}\|^2_{\mathcal{M}} + \frac{\alpha}{2} \|\nabla \boldsymbol{u}\|^2_{\mathcal{Q}} \right\}, \quad \nabla \cdot \boldsymbol{u} = 0, \tag{14}$$

where α is a tunable parameter. Relying on a Lagrangian formalism as exposed in §2.2, an adjoint variable π, which may be physically interpreted as a pseudo-pressure field, is introduced in association with the divergence-free condition. As (14) forms a quadratic problem, its solution may be directly obtained through inverting the following linear system

$$\begin{pmatrix} -\alpha\Delta \circ + \mathbf{H}^\dagger \mathbf{H} \circ & \nabla \circ \\ -\nabla \cdot \circ & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{u} \\ \pi \end{pmatrix} = \begin{pmatrix} \mathbf{H}^\dagger \boldsymbol{m} \\ 0 \end{pmatrix}. \tag{15}$$

The Neumann condition $\nabla \boldsymbol{u} \cdot \boldsymbol{n} = \mathbf{0}$, where $\boldsymbol{n}$ refers to the normal unit vector to a boundary, is imposed by default at boundaries, unless Dirichlet boundary conditions are available. An appropriate value for the regularization parameter α is obtained through the L-curve technique [36]. While such an approach requires several inversions of (15) (typically ~ 10), we emphasize that the associated computational cost remains negligible compared to that of the whole optimization procedure in §2.2. Furthermore, it was verified that finding a specific optimum value of α was not a strong necessity, as reasonable variations in the smoothness of the first-guess flow do not appear to have a significant influence on the reconstructed flow.

Secondly, a first guess for the pressure field p is evaluated. For this purpose, we can not directly rely on the pseudo-pressure π that is given through solving (15). This field is indeed highly unphysical (pointwise), which may be inferred through considering that π is formally solution to the Poisson problem $\Delta\pi = -\nabla \cdot (\mathbf{H}^\dagger(\mathbf{H}\boldsymbol{u} - \boldsymbol{m}))$, whose irregular right-hand-side is given through (11). Instead, we solve the Poisson problem

for pressure $\Delta p = -\nabla\cdot((\boldsymbol{u}\cdot\nabla)\boldsymbol{u})$ based on the velocity field obtained from the previous step. The determination of appropriate boundary conditions for this Poisson problem is known to be a delicate task [4], and as we assume to solely rely on the measurements $\boldsymbol{m}$, $\nabla p\cdot\boldsymbol{n}=0$ is imposed at all boundaries for the sake of simplicity, the pressure field being automatically corrected during the data assimilation procedure.

Thirdly, based on the velocity and pressure fields from previous steps, the left-hand-side of the momentum equation in (4)-(5) may be evaluated, which straightforwardly provides a first guess for the forcing $\boldsymbol{f}$.

In the following, the output of the above procedure will be referred to as the first-guess flow.

3.2. Gradient regularization

As scattered pointwise/PTV measurements are here considered, regularization has to be considered in the data assimilation procedure in order not to introduce spurious discontinuities linked to the introduction of such pointwise information. A first known strategy to address this aspect may be penalization (see, e.g., [35, 32]), namely adding terms in the cost function J to minimize in (6) that would penalize the spatial gradient (or higher-order derivatives) of the control vector $\boldsymbol{f}$. As an alternative, we here consider H^1-like regularization of the gradient $\frac{dJ}{d\boldsymbol{f}}$ in (12) [28, 29, 30]. From such a gradient, a regularized gradient $\frac{dJ^{H^1}}{d\boldsymbol{f}}$ is obtained based on the definition

$$\int_\Omega \frac{dJ}{d\boldsymbol{f}}\cdot\tilde{\boldsymbol{f}}\,d\Omega = \frac{1}{1+l_{\text{sob}}^2}\int_\Omega\left\{\frac{dJ^{H^1}}{d\boldsymbol{f}}\cdot\tilde{\boldsymbol{f}}+l_{\text{sob}}^2\nabla\left(\frac{dJ^{H^1}}{d\boldsymbol{f}}\right)^{\mathrm{T}}:\nabla\tilde{\boldsymbol{f}}\right\}d\Omega, \quad (16)$$

for all admissible $\tilde{\boldsymbol{f}}$. While the left-hand-side of (16) corresponds to a L^2 scalar product (which also corresponds to the definition of $\langle\circ,\circ\rangle_{\mathcal{Q}}$ in §2.2), its right-hand-side is close to the definition of a H^1 scalar product, where the parameter l_{sob} allows to adjust the degree of smoothness of $\frac{dJ^{H^1}}{d\boldsymbol{f}}$. When getting $\frac{dJ^{H^1}}{d\boldsymbol{f}}$ based on (16), one may here want to preserve the divergence-free character of the gradient, as $\boldsymbol{f}$ should account for a divergence-free quantity (see §2.4). Accordingly, verifying (16) in conjunction with the imposition of $\nabla\cdot\frac{dJ^{H^1}}{d\boldsymbol{f}}=0$ may be formulated as a standard equality-constrained quadratic-programming problem not far from that in (14), and whose solution is provided through the inversion of

$$\begin{pmatrix}\frac{1}{1+l_{\text{sob}}^2}(\mathbb{I}-l_{\text{sob}}^2\Delta)\circ & \nabla\circ\\ -\nabla\cdot\circ & 0\end{pmatrix}\begin{pmatrix}\frac{dJ^{H^1}}{d\boldsymbol{f}}\\ \pi\end{pmatrix}=\begin{pmatrix}\frac{dJ}{d\boldsymbol{f}}\\ 0\end{pmatrix}, \quad (17)$$

where, similarly as in (15), π is introduced to take into account the divergence-free constraint. Boundary conditions are the same as for the adjoint velocity field $\boldsymbol{u}^\dagger$ when solving the adjoint problem (9) (in virtue of (12), see also §2.3). It may be worth noting that $\frac{dJ^{H^1}}{d\boldsymbol{f}}$ remains a descent direction, as demonstrated in, e.g., [29], and is thus fully adapted to solve the constrained optimization problem of our NS-DA-SIM method.

Although such a regularization approach requires solving the supplementary problem (17), one may identify at least two advantages compared to penalization techniques. Firstly, the use of (17) ensures that the control vector is always updated in a sufficiently smooth way at every iteration of the optimization procedure, thus favouring the stability of the computations. This is not the case in the penalization approach, which only enables the control vector to be smoothed out over the whole optimization procedure. Secondly, the parameter l_{sob} in (16)-(17) may be interpreted as a (non-dimensional) filter length [28]. This allows to choose l_{sob} based on physical arguments, while parameters in penalization approaches are less straightforward to interpret and tune, and usually require multiple optimizations to identify appropriate values. In the present case, as gradient regularization is employed in order to tackle the sparsity of the measurements, we choose $l_{\text{sob}} = \overline{d}$, where $\overline{d}$ is the mean distance between measurements, as introduced in §2.1. The validity of this choice will be confirmed in §4-§5 through the physically-sound character of the reconstruction results.

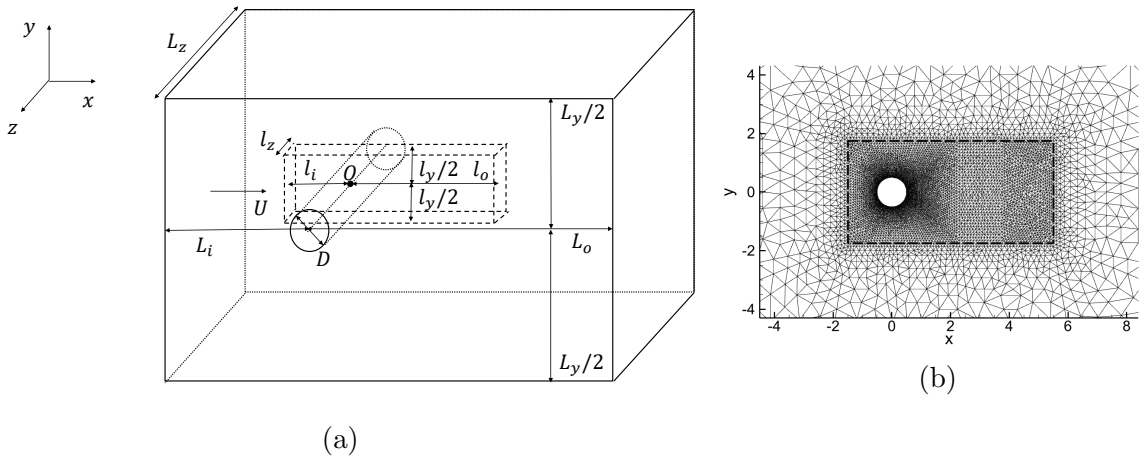
3.3. Discretization methods

Spatial discretization is performed based on the finite-element method as implemented in the software FreeFEM++ [37]. To allow affordable computations in the NS-DA-SIM methodology, piecewise-linear functions that are enriched by bubble functions are used for the velocity space, while piecewise-linear functions are used for pressure [38]. Second-order polynomials elements for the velocity space are exceptionally used when generating the numerical reference flow in §4. Steady linear problems such as (9), (15) and (17) are solved in a fully coupled way through a direct matrix inversion. The nonlinearity of the steady Navier-Stokes equations (4) is handled through the Newton method. Still for (4), in order to facilitate the consideration of relatively high-Reynolds number flows such as in §5, streamline-upwind Petrov–Galerkin (SUPG) [39, 35] and grad-div [40] stabilisations, along with pseudo-time stepping [41], are implemented. When considered unsteady, in particular to generate the numerical reference flow in §4, the Navier-Stokes equations (2) are solved based on a fractional step methodology [42]. Time integration is performed in a fully implicit way based on a second-order finite-difference approximation of the time derivative. Finally, while the adjoint problem (9) has been derived following the continuous approach in §2.2 for the sake of simplicity, it is actually implemented following the discrete adjoint approach [31]. The same numerical methods are employed for the alternative data assimilation methodologies that are described in Appendix A.

4. Numerical case: flow around a cylinder at $Re = 300$

4.1. Case setup

The present NS-DA-SIM methodology is here applied to the reconstruction of the numerical flow past a cylinder at Reynolds number $Re = 300$, based on the diameter of the cylinder D and the uniform inlet velocity U, which are used to non-dimensionalize



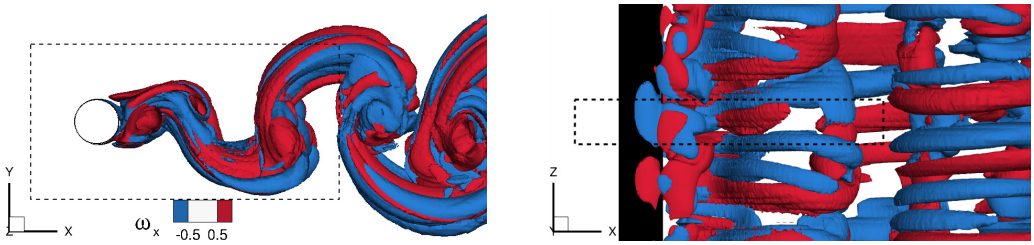
(a) (b)

Figure 1: (a) Schematic of the computational domain for the 3D flow past a circular cylinder at $Re = 300$. (b) 2D slice of the mesh, zoomed in the vicinity of the cylinder. In both figures, dashed lines delineate the domain in which synthetic measurements are extracted.

all quantities in this section. The flow configuration and computational domain are depicted in figure 1a. Coordinates x, y and z denote the streamwise, cross-stream and spanwise directions, respectively. The origin O of the frame lies at the center of the cylinder, which corresponds to the mid point of its axis.

A reference direct numerical simulation (DNS) is first performed, where the distance of the cylinder axis to the inlet, outlet and top and bottom boundaries are $L_i = 20$, $L_o = 30$ and $L_y/2 = 20$, respectively, while the extent of the computational domain in the spanwise direction is $L_z = 12$, as in the DNS study in [43]. The uniform boundary condition $\boldsymbol{u} = (u, v, w)^{\mathrm{T}} = (1, 0, 0)^{\mathrm{T}}$ is imposed at the inlet, $(Re^{-1}\frac{\partial u}{\partial x} - p, \frac{\partial v}{\partial x}, \frac{\partial w}{\partial x}) = \mathbf{0}$ is used at the outlet, symmetry conditions $(\frac{\partial u}{\partial y}, v, \frac{\partial w}{\partial y}) = \mathbf{0}$ are imposed at top and bottom boundaries, the no-slip condition $\boldsymbol{u} = \mathbf{0}$ is enforced at the cylinder, while periodic boundary conditions are used at the transverse planes $z = \pm L_z/2$. The 3D mesh is obtained by extruding a 2D mesh in the transverse direction which is discretized with 90 planes, resulting in a 3D mesh that is formed of tetrahedra with a total of $3.9 \cdot 10^6$ vertices. The results obtained with the present discretization are very close to those obtained in [43]; in particular, we here estimate the Strouhal number and average drag coefficient over 500 convective times as $St = 0.204$ and $\overline{C_d} = 1.28$, which correspond to discrepancies of the order of 0.1% and less than 1%, respectively, with the highest-resolution DNS results of [43].

The instantaneous flow is illustrated in figure 2 where isosurfaces of the streamwise vorticity $\omega_x = \frac{\partial w}{\partial y} - \frac{\partial v}{\partial z}$ are reported. This quantity enables to assess purely 3D effects and the identification of so-called mode B structures, whose spanwise wavelength is usually found between 0.8 and 1 [43, 44, 45]. Synthetic observations are extracted from this instantaneous flow in the domain that is delineated by dashed lines in figures 1-2. The distances of the cylinder axis to the boundaries of this measurement domain, as



(a) (b)

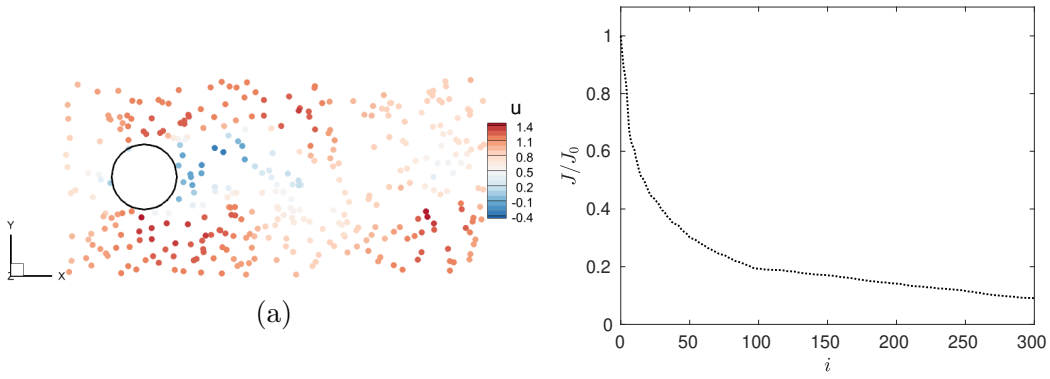
Figure 2: Isosurfaces of streamwise vorticity $\omega_x = \pm 0.5$ for the reference instantaneous DNS flow in (a) z-normal and (b) y-normal views. Dashed lines delineate the domain in which synthetic measurements are extracted.

depicted in figure 1a, are $l_i = 1.5$, $l_o = 5.5$ and $l_y/2 = 1.75$, while the extent in the spanwise direction is $l_z = 1$ in order to include at least one whole mode B structure. The present measurement domain thus extends both upstream and downstream of the cylinder, which is not necessarily typical of actual experiments [46, 20, 24]. However, this choice will allow us to illustrate the ability of the present NS-DA-SIM methodology in reconstructing high-gradient regions and in providing aerodynamic efforts. This setup will be rediscussed in the following. Measurement locations $\boldsymbol{x}_i$ are randomly generated in the measurement domain to match a prescribed mean distance $\overline{d}$, which will be varied in the following. Synthetic measurements $\boldsymbol{m}$ then correspond to the value of the DNS velocity field at these locations. In the following, the aim of the NS-DA-SIM methodology will thus be the reconstruction of the full instantaneous flow of figure 2 in the measurement domain from these observations.

The computational domain that is employed for the NS-DA-SIM procedure is similar to that for the above DNS, but restricting its spanwise extent to that of the measurement domain, i.e. taking $L_z = l_z = 1$. The corresponding 2D mesh that is extruded in the spanwise direction to get the 3D mesh is illustrated in figure 1b. It is mainly refined in the measurement domain, especially close to the cylinder, with a density of vertices nearly identical to that of the mesh for the reference DNS. The resulting 3D mesh is formed of tetrahedra with a total of $8.7 \cdot 10^4$ vertices. Imposed boundary conditions are similar as described above, except at the transverse planes $z = \pm L_z/2 = \pm 0.5$ where the first-guess velocity field is imposed as Dirichlet condition (see §2.3). This computational domain along with the measurement one will be varied and restricted downstream of the cylinder to get closer to experimental conditions in §4.4.

4.2. Reconstruction results

We first consider relying on relatively dense synthetic measurements $\boldsymbol{m}$ with mean distance $\overline{d} = 0.08$. This value coincides with one tenth of the lower-bound estimation of the spanwise wavelength of mode B structures. The locations of these synthetic



(a) (b)

Figure 3: (a) Synthetic measurement locations in z-normal view for $\overline{d} = 0.08$ colored by the streamwise velocity u of the reference flow, in a $-0.025 \leq z \leq 0.025$ slice. (b) Cost function J (normalized by its initial value J_0) versus the iteration of the minimization procedure in the NS-DA-SIM approach.

measurements are illustrated in figure 3a in z-normal view, restricting the shown locations to $-0.025 \leq z \leq 0.025$. The convergence of the minimization procedure in the NS-DA-SIM methodology is illustrated in figure 3b. It appears that the value of the cost function J in (6) has been decreased by more than one order of magnitude in 300 iterations. While this number of iterations has been considered in the present case, as in the following ones, to ensure converged results, already satisfactory solutions could be obtained through only ~ 100 iterations. Figures 4-6 illustrate the estimation of various flow quantities, comparing the reconstructed flow with the reference and first-guess ones. In addition, figure 7 reports the following reconstruction relative error for various quantities b

$$e(b) = \frac{\int_{\Omega_m} (b^r - b)^2 \, d\Omega_m}{\int_{\Omega_m} b^{r\,2} \, d\Omega_m}, \tag{18}$$

where b^r and b refer to the reference and reconstructed fields, respectively, and Ω_m is the measurement domain. Errors are reported not only for the present case $\overline{d} = 0.08$, which is highlighted through vertical grey dotted lines, but also for other measurement spatial samplings which will be discussed in the following. While this mean inter-particle distance is here given in non-dimensional form (based on the cylinder diameter D), it should be emphasized that the chosen values are fully in line with typical seeding densities obtained in experiments. In particular, in the case of the jet experiment in §5, a seeding density of 0.06 particles per pixel (ppp) corresponds to $\overline{d} = 0.06$, and 0.02 ppp to $\overline{d} = 0.09$.

It appears from figure 7a that the estimation of the velocity field for the first-guess flow is already quite satisfactory, with $e(\boldsymbol{u}) = 4.5 \cdot 10^{-3}$. This is due to the relatively high density of the measurements with respect to flow structures of size ~ 1 in this case. Still, this estimation is significantly further improved by the NS-DA-SIM

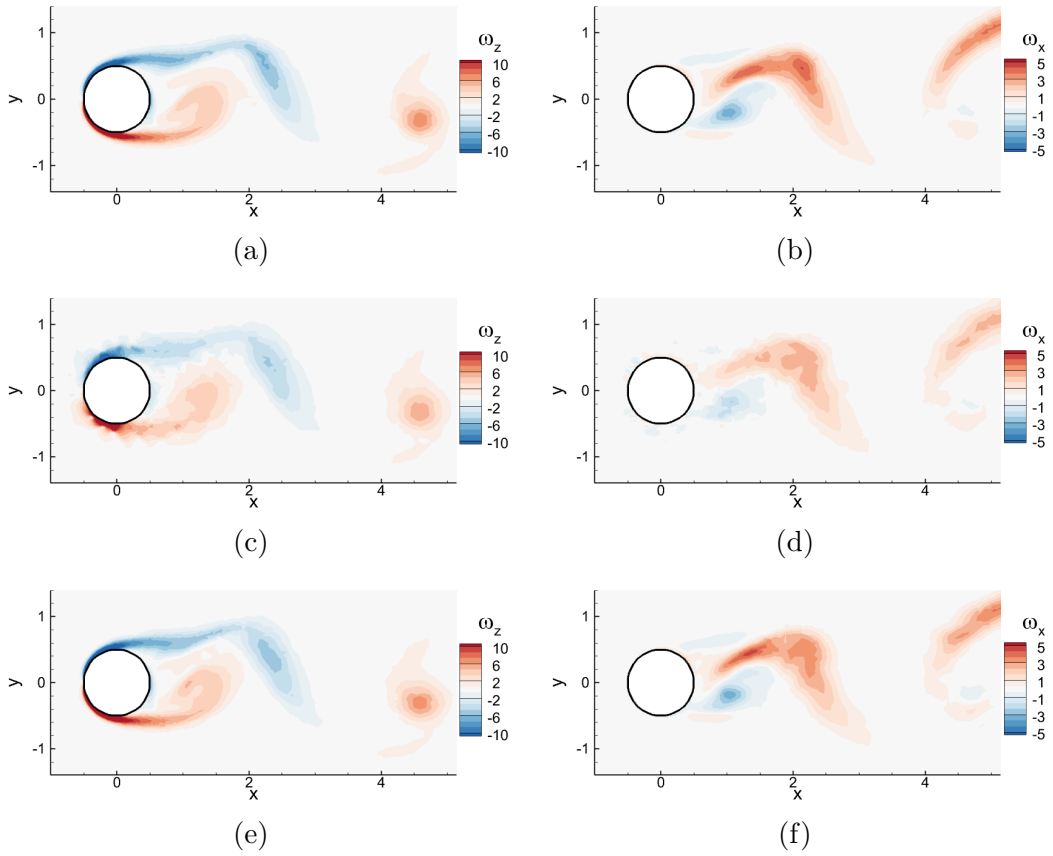


Figure 4: Data assimilation with synthetic measurements corresponding to $\overline{d} = 0.08$. Left and right columns report the vorticity components ω_z and ω_x in the $z = 0$ plane, respectively. Top to bottom rows report results for (a,b) the reference flow, (c,d) the first-guess flow, and (e,f) the reconstructed flow from the NS-DA-SIM approach, respectively.

approach ($e(\boldsymbol{u}) = 7.7 \cdot 10^{-4}$). We directly illustrate the more challenging estimation of the velocity gradients through the vorticity components $\omega_z = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}$ and $\omega_x = \frac{\partial w}{\partial y} - \frac{\partial v}{\partial z}$ in figures 4 and 5. The reconstruction of the vorticity component ω_z in the $z = 0$ plane is first considered through figures 4a, 4c and 4e. While the 2D features of the vortex shedding are overall captured in the first-guess flow, the latter provides a crude estimation of ω_z close to the cylinder, in particular regarding the shear layers at the top and bottom of the cylinder. The NS-DA-SIM approach enables a large enhancement in the reconstructed ω_z field, which appears very close to the true one. The reconstruction of purely 3D flow features as quantified through the vorticity component ω_x is illustrated in figures 4b, 4d and 4f in the $z = 0$ plane, and in figure 5 in the $y = 0.25$ and $y = 0.5$ planes. The reference ω_x field appears accurately captured in the reconstructed flow, which is not as much the case for the first-guess one. Overall, as reported by figure 7b, the NS-DA-SIM methodology has decreased by almost one order of magnitude the error $e(\nabla \boldsymbol{u})$ on the velocity gradients compared to the first-guess flow, from $e(\nabla \boldsymbol{u}) = 1.9 \cdot 10^{-1}$

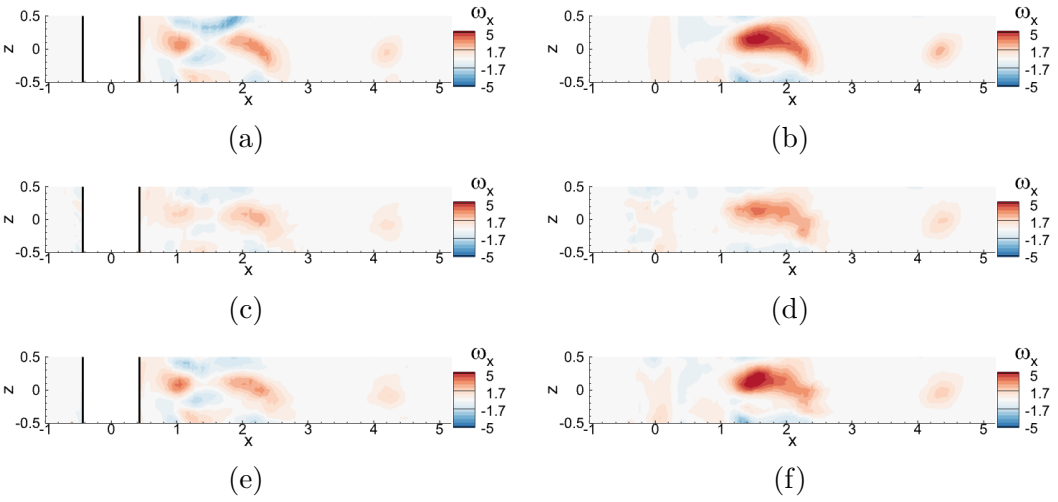


Figure 5: Data assimilation with synthetic measurements corresponding to $\overline{d} = 0.08$. Left and right columns report the streamwise vorticity ω_x in the $y = 0.25$ and $y = 0.5$ planes, respectively. Top to bottom rows report results for (a,b) the reference flow, (c,d) the first-guess flow, and (e,f) the reconstructed flow from the NS-DA-SIM approach, respectively.

to $e(\nabla \boldsymbol{u}) = 3.0 \cdot 10^{-2}$.

The ability of the NS-DA-SIM approach in inferring other quantities such as the pressure field, but also Eulerian acceleration, is investigated in figure 6 for the $z = 0$ plane. First concerning pressure (see figures 6a, 6c and 6e), the first-guess flow appears to both underestimate the reference pressure upstream of the cylinder and overestimate it in the wake. In contrast, the reconstructed pressure field appears very satisfactory and close to the true one, with $e(p) = 9.8 \cdot 10^{-3}$ (while $e(p) = 5.3 \cdot 10^{-2}$ for the first-guess flow). A valuable outcome of this accurate reconstruction of the pressure field along with that of the velocity gradients is the good estimation of the instantaneous aerodynamic coefficients with the NS-DA-SIM approach, as reported in figures 7e-7f, with remaining discrepancies that are below 3% for both drag and lift coefficients.

The y-component of the Eulerian acceleration field $\frac{\partial v}{\partial t}$, which is particularly considered as it takes the highest values in magnitude among the three components, is illustrated in figures 6b, 6d and 6f. This field is thus estimated from the forcing $\boldsymbol{f}$ according to $\frac{\partial v}{\partial t} = -f_y$ for the first-guess and reconstructed flows. The intensity of $\frac{\partial v}{\partial t}$ appears significantly underestimated for the first-guess flow compared to the reference one ($e(\frac{\partial \boldsymbol{u}}{\partial t}) = 1.8 \cdot 10^{-1}$). The estimation of the reconstructed flow is more satisfactory ($e(\frac{\partial \boldsymbol{u}}{\partial t}) = 9.2 \cdot 10^{-2}$), even if remaining discrepancies are more visible than for other quantities. The inspection of figures 7a-7d confirms that Eulerian acceleration is the most difficult quantity to accurately reconstruct among the considered ones. The present results appear reminiscent of the findings in [32], where it was noted in the context of mean-flow reconstruction that significantly fewer measurements were needed to ensure a good estimation of the full mean velocity field than of the control vector of the data

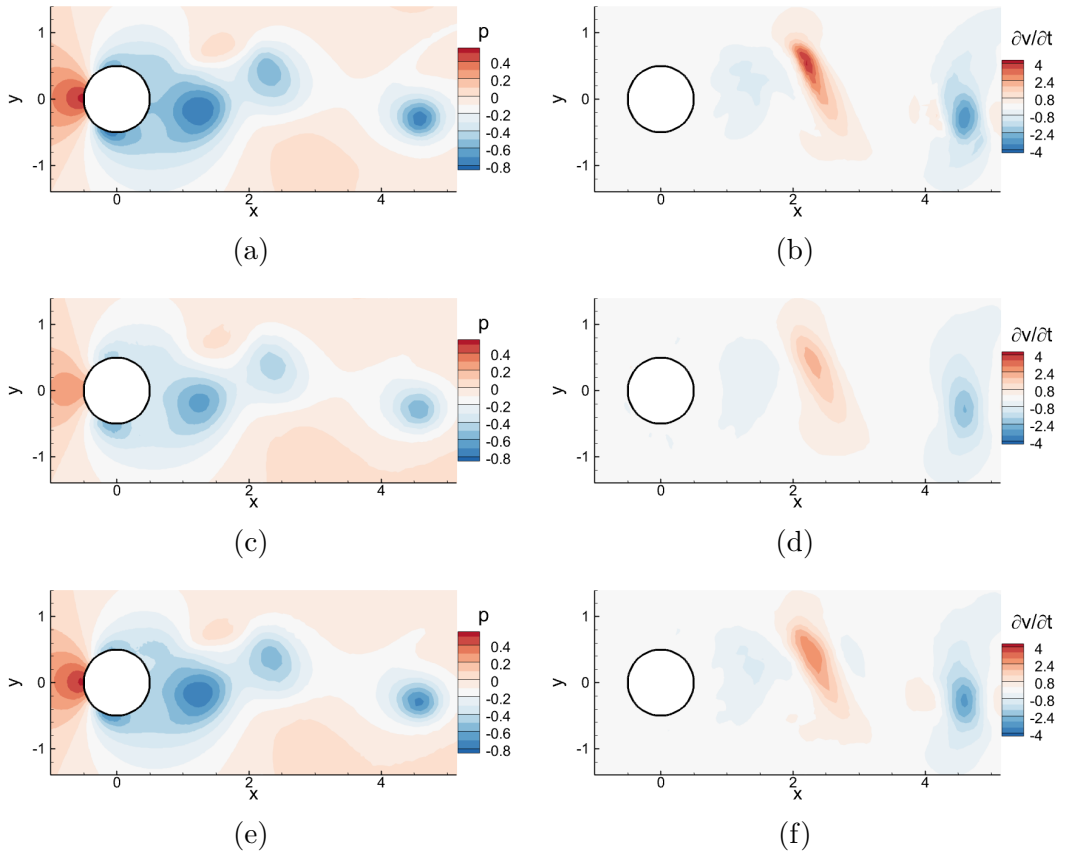


Figure 6: Data assimilation with synthetic measurements corresponding to $\overline{d} = 0.08$. Left and right columns report the pressure field p and the Eulerian acceleration component $\frac{\partial v}{\partial t}$ in the $z = 0$ plane, respectively. Top to bottom rows report results for (a,b) the reference flow, (c,d) the first-guess flow, and (e,f) the reconstructed flow from the NS-DA-SIM approach, respectively.

assimilation procedure, which was also a forcing in the momentum equation [26]. As indicated by the above discussed results, the noted discrepancies in the inferred forcing $\boldsymbol{f}$/ Eulerian acceleration do not however preclude an accurate estimation of other flow quantities of interest, while figure 7d confirms that the identification of the Eulerian acceleration field may be further improved in the case of denser measurements.

Figure 7 also confirms the abilities of the present data assimilation approach in flow reconstruction based on fewer PTV measurements. For all considered values of $\overline{d}$, it provides significant improvements compared to the first-guess flow, even for the highest values of $\overline{d}$, which correspond to very low measurement densities.

4.3. Comparisons with other data assimilation approaches

The present NS-DA-SIM methodology is now compared with other variational data assimilation approaches that are based on alternative physical models and detailed in Appendix A. In order to highlight the importance of the physical model in the fidelity

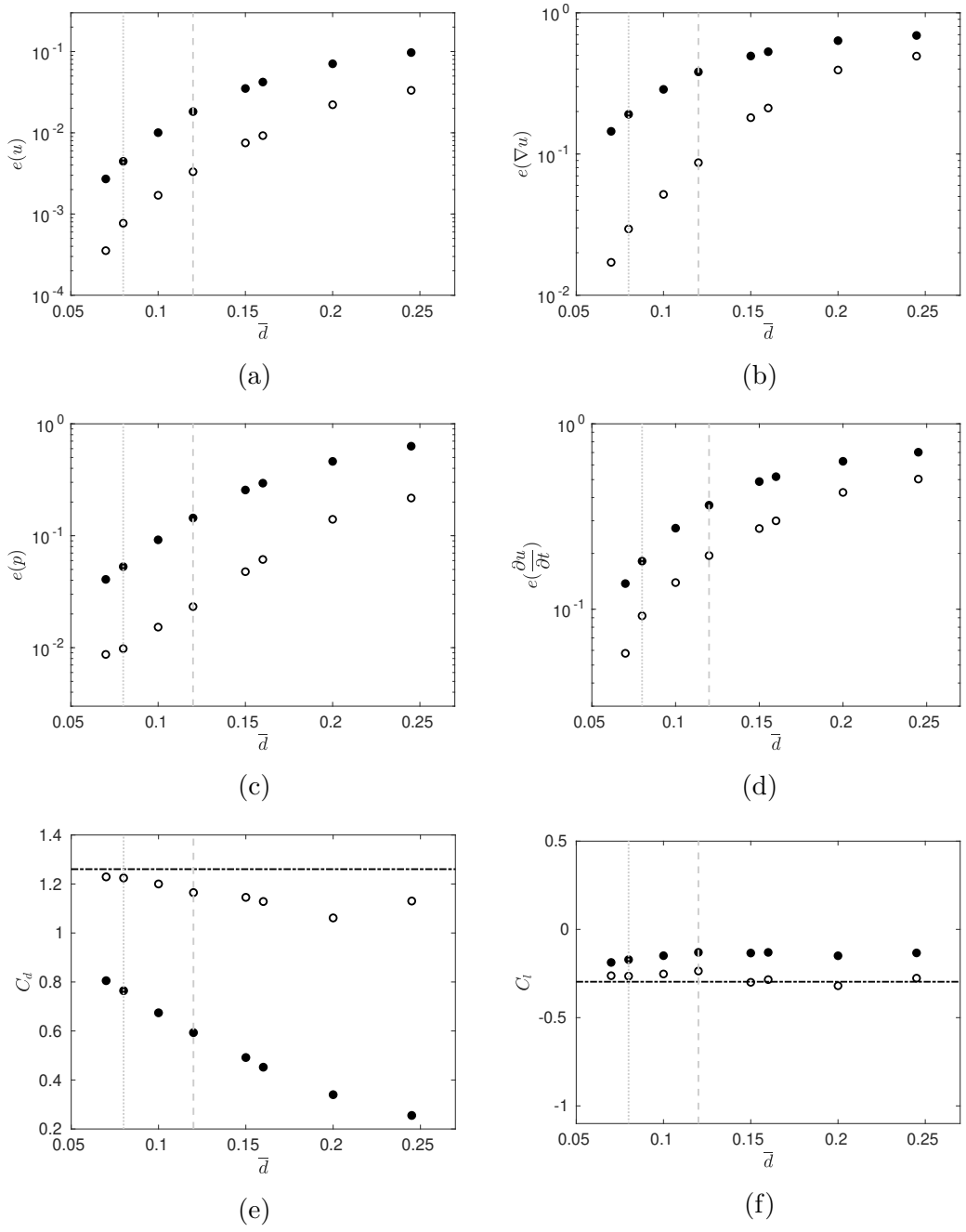


Figure 7: (a-d) Relative reconstruction error $e(b)$ (defined in (18)) with b taken as (a) the velocity field $\boldsymbol{u}$, (b) the velocity gradient $\nabla \boldsymbol{u}$, (c) the pressure field p, and (d) the Eulerian acceleration $\frac{\partial \boldsymbol{u}}{\partial t}$, versus the mean distance between synthetic measurements $\overline{d}$. (e-f) Estimated (e) drag C_d and (f) lift C_l coefficients, reference values are reported through horizontal black dash-dotted lines. Results are reported for the first-guess flow (●) and the reconstructed flow from the NS-DA-SIM approach (○). Vertical grey dotted and dashed lines indicate $\overline{d} = 0.08$ and $\overline{d} = 0.12$, respectively.

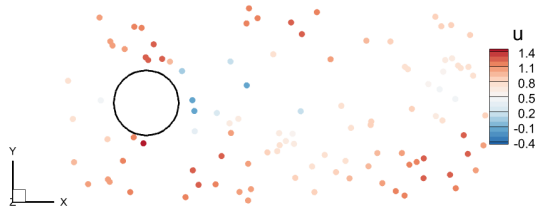


Figure 8: Synthetic measurement locations for $\overline{d} = 0.12$ in z-normal view colored by the streamwise velocity u of the reference flow, in a $-0.025 \leq z \leq 0.025$ slice.

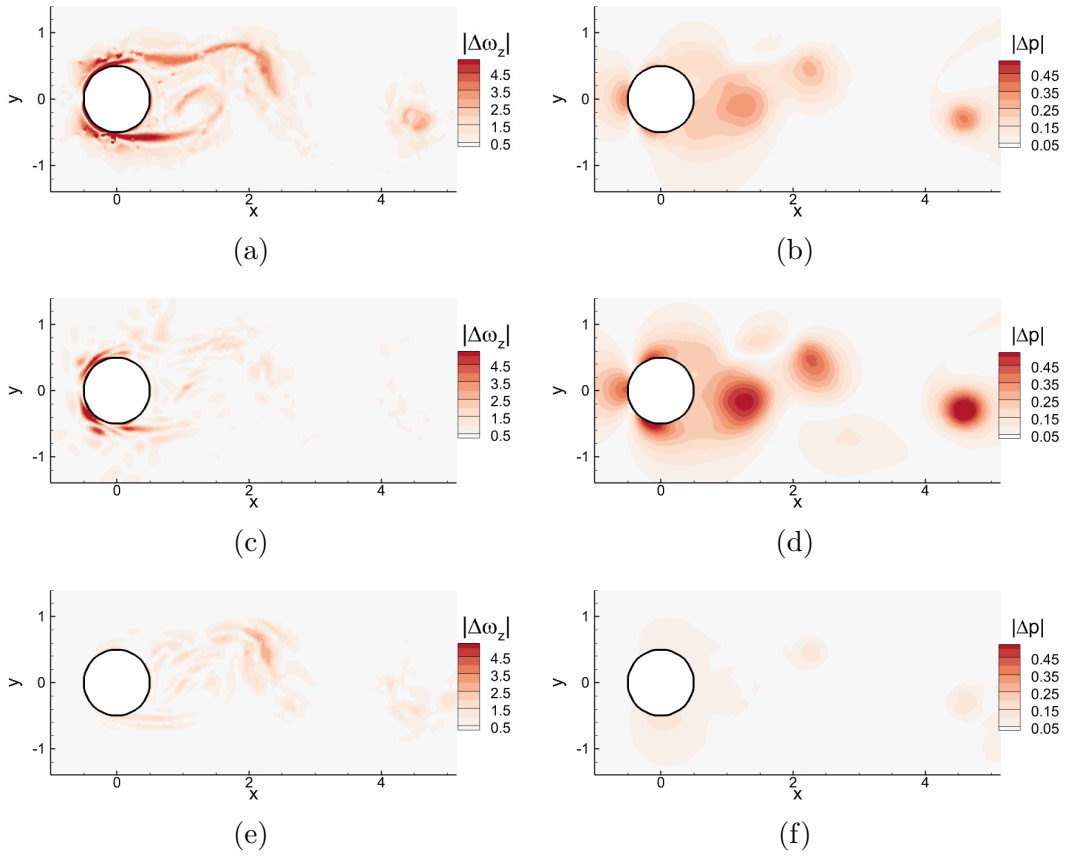


Figure 9: Data assimilation with synthetic measurements corresponding to $\overline{d} = 0.12$: comparisons with Stokes-based data assimilation. Left and right columns report the absolute error field between reference and estimated quantities $|\Delta b| = |b^r - b|$ in the $z = 0$ plane with b taken as the vorticity component ω_z or the pressure field p, respectively. Top to bottom rows report results for (a,b) the first-guess flow, (c,d) the reconstructed flow from the Stokes-based approach, and (e,f) the reconstructed flow from the NS-DA-SIM approach, respectively.

of the reconstruction, comparisons are performed based on a sparser set of observations than in previous figures 3-6, namely $\overline{d} = 0.12$ as illustrated in figure 8. As $\overline{d} \propto N_{\mathrm{PTV}}^{-\frac{1}{3}}$, the present case is indeed roughly equivalent to three times less data than illustrated in figure 3a ($\overline{d} = 0.08$).

We first investigate the possibility of relying on the strong imposition of the simpler and linear Stokes equations in the data assimilation procedure, similarly as performed in [47] (however working directly on particle image data), instead of the (steady) Navier-Stokes equations in the present NS-DA-SIM methodology. In this case, as detailed in Appendix A.1, the control vector of the minimization procedure represents Lagrangian acceleration, namely $\frac{d\boldsymbol{u}}{dt} = \frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{u} \cdot \nabla)\boldsymbol{u}$. The same first-guess flow is used for both this so-called Stokes-based approach and the present NS-DA-SIM methodology (see §3.1 and Appendix A.1).

Reconstruction results are illustrated in figure 9, where the absolute discrepancies between the reference and reconstructed vorticity component ω_z and pressure p in the $z = 0$ plane are reported. The Stokes-based approach allows already a significant improvement in the estimation of the velocity field and in particular of the vorticity component ω_z compared to the first-guess flow. On the other hand, remaining discrepancies close to the cylinder, and in particular upstream, are still large. This may suggest limitations of the Stokes-based approach in correctly accounting for relatively high velocity gradients, in particular close to solid boundaries. Furthermore, it appears that the discrepancies with respect to the reference pressure field have been largely increased compared to the first-guess flow. This is explained by the fact that only the divergence-free part of Lagrangian acceleration may be identified through the Stokes-based data assimilation procedure, leading to an indeterminacy in the pressure estimation (see §2.4 and Appendix A.1). The present results thus confirm the inability of Stokes-based methods in reconstructing pressure. In contrast, the present NS-DA-SIM approach succeeds in decreasing by almost one order of magnitude the errors in the pressure estimation compared to the first-guess flow (see also figure 7c). In addition, concerning the reconstruction of the vorticity component ω_z, the present approach enables an accurate reconstruction of the flow close the cylinder and, despite of slightly larger remaining discrepancies in the wake, provides an overall better estimation of the velocity gradients compared to the Stokes-based approach ($e(\nabla \boldsymbol{u}) = 1.3 \cdot 10^{-1}$ and $e(\nabla \boldsymbol{u}) = 8.7 \cdot 10^{-2}$ for the Stokes-based and NS-DA-SIM procedures, respectively). These findings therefore confirm the advantage of solving for the convection term in the present approach, in contrast to Stokes-based methodologies, for the estimation of both velocity and pressure fields.

Data assimilation based on the unsteady Navier-Stokes equations (see, e.g., [20, 25, 24]), which will be hereafter referred to as 4DVar (following a general terminology, see, e.g., [9]), is now briefly considered. In the present context of single-instant measurements, 4DVar may be applied as follows: the measurements $\boldsymbol{m}$ and the flow to reconstruct are assumed to correspond to the end of a time window of size T, and the control vector of the data assimilation procedure is the initial flow at the beginning

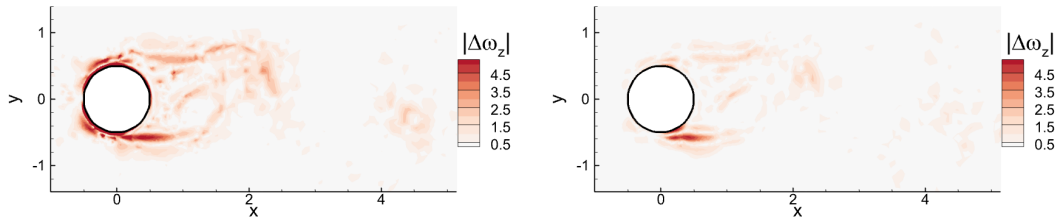


Figure 10: Unsteady data assimilation (4DVar) with synthetic measurements corresponding to $\overline{d} = 0.12$. The error $|\Delta\omega_z| = |\omega_{z,r} - \omega_z|$ in the $z = 0$ plane is reported for time window sizes (a) $T = 0.025$ and (b) $T = 0.25$.

of this time window. More formulation and implementation details may be found in Appendix A.2. In this approach, the time window size T is an adjustable parameter. Reconstruction results obtained with 4DVar are illustrated in figure 10 for $T = 0.025$ and $T = 0.25$. While one could be tempted to use a short time horizon in this context of single-instant measurements, it clearly appears that 4DVar is not able to provide accurate reconstruction results for the case $T = 0.025$, for which there is only moderate improvement compared to the first-guess flow (see figure 9a). This suggests that T has to be large enough to enable the different terms in the unsteady Navier-Stokes equations to act between measurement locations and significantly improve the first-guess flow. While the determination of an optimal value for T and of its dependency with respect to the measurement density and other physical parameters is out of the scope of this study, it is found that increasing it to $T = 0.25$ in the present case allows to obtain reconstruction results whose quality becomes similar to that of the estimation from the present NS-DA-SIM methodology ($e(\nabla \boldsymbol{u}) = 1.0 \cdot 10^{-1}$ and $e(\nabla \boldsymbol{u}) = 8.7 \cdot 10^{-2}$ for the 4DVar (with $T = 0.25$) and NS-DA-SIM procedures, respectively). The non-trivial determination of an appropriate time horizon for the consideration of single-instant measurements, in conjunction with the inherent complexity in implementing 4DVar and performing it over appreciable time horizons - which necessitates a large quantity of memory or disk storage associated to the direct flow snapshots within the time horizon - emphasize in our view the benefits of the proposed approach in the present context. This may be further supported by the fact that one may expect these difficulties in 4DVar to be significantly amplified for flows at higher Reynolds number, all the more as their smaller Lyapunov time scales are known to require elaborate strategies to make 4DVar converge [23].

4.4. Variations in the measurement and computational domains

Finally in this section, we consider the case where measurements are restricted to a region downstream of the cylinder, which may be more representative of actual experimental setups [46, 20, 24]. Observations are supposed to be available only from

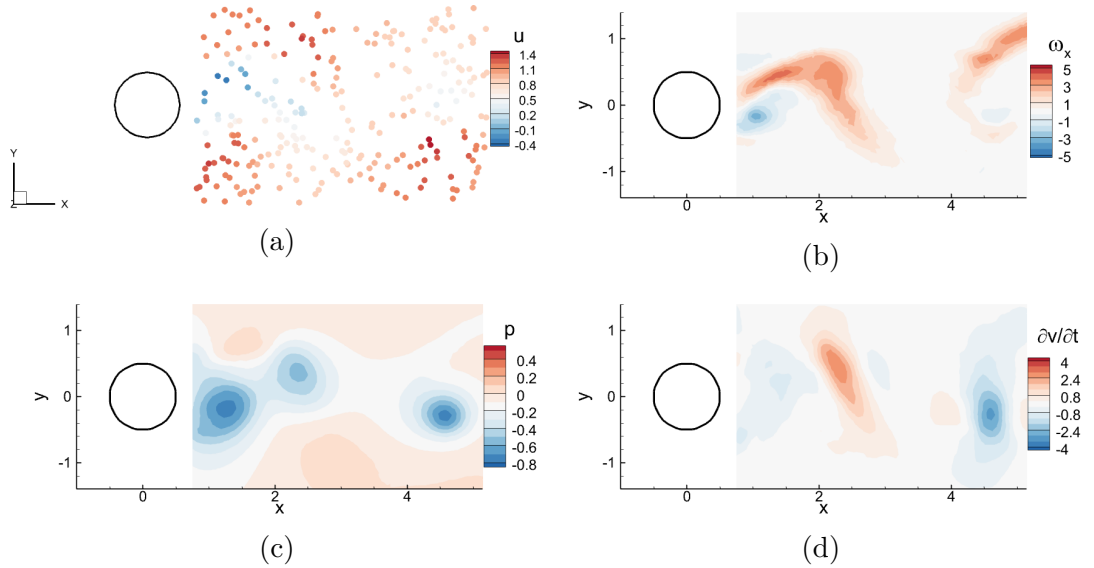
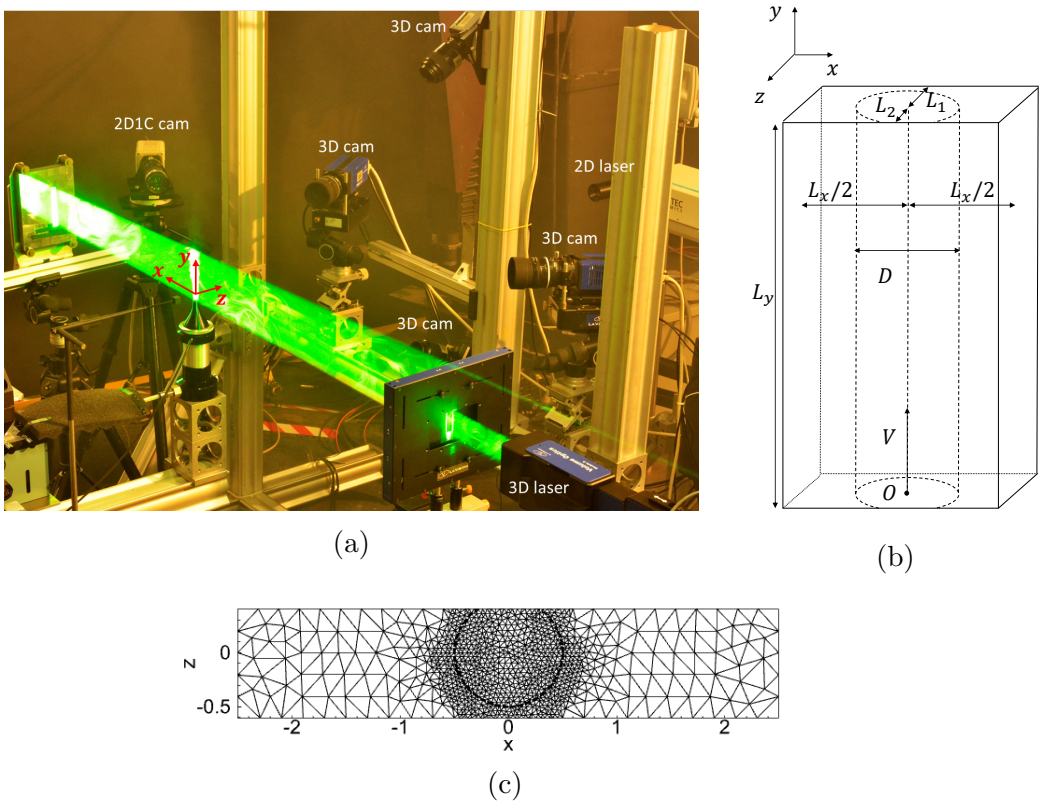


Figure 11: Data assimilation where both the synthetic measurements and the computational domain are restricted to a region downstream of the cylinder. (a) Synthetic measurement locations for $\overline{d} = 0.08$ in z-normal view colored by the streamwise velocity u of the reference flow, in a $-0.025 \leq z \leq 0.025$ slice. (b-d) Recontructed (b) streamwise vorticity ω_x, (c) pressure p and (d) Eulerian acceleration component $\frac{\partial v}{\partial t}$ in the $z = 0$ plane with the NS-DA-SIM approach.

$x = 0.75$, while the other dimensions of the measurement domain are the same as previously (see §4.1). The computational domain is chosen to fully coincide with the measurement domain, which allows to minimize computational costs. The inlet condition thus becomes non-trivial and non-uniform in this case as the inlet intersects the wake flow in transverse directions, and is now determined based on the first-guess velocity field (see §2.3). The considered observations, for which $\overline{d} = 0.08$ as for most of the reported results in §4.2, along with the reconstructed flow with the NS-DA-SIM approach are illustrated in figure 11. The reconstructed flow appears similar to the one in figures 4 and 6 obtained with the previous measurement and computational domains. This is confirmed by the similar reconstruction errors evaluated in the present measurement/computational domain: $e(\boldsymbol{u}) = 7.8 \cdot 10^{-4}$ (respectively $e(\boldsymbol{u}) = 8.7 \cdot 10^{-4}$), $e(\nabla \boldsymbol{u}) = 5.3 \cdot 10^{-2}$ (respectively $e(\nabla \boldsymbol{u}) = 5.8 \cdot 10^{-2}$), $e(\boldsymbol{p}) = 1.7 \cdot 10^{-2}$ (respectively $e(\boldsymbol{p}) = 1.2 \cdot 10^{-2}$) and $e(\frac{\partial \boldsymbol{u}}{\partial t}) = 7.6 \cdot 10^{-2}$ (respectively $e(\frac{\partial \boldsymbol{u}}{\partial t}) = 9.3 \cdot 10^{-2}$) in the present case (respectively when relying on the previous measurement and computational domains). These findings demonstrate the robustness of the present NS-DA-SIM procedure with respect to the choice of the computational domain, which may be chosen to fully coincide with the measurement one even if its boundaries (even the inlet one) cut fluctuating parts of the flow. This choice will be adopted for the following experimental test case in §5.



(a) (b)

(c)

Figure 12: (a) Experimental setup of the cylindrical transitional jet at $Re = 4600$, including volumetric illumination and four cameras for 3D PTV measurements. Reprinted by permission from Springer Nature [3]. (b) Schematic of the computational domain for data assimilation, which identifies with the measurement domain. (c) 2D slice of the mesh used for the computations. Dashed lines in (b,c) symbolize the jet core.

5. Reconstruction of an experimental round jet flow at $Re = 4600$

5.1. Experimental and numerical setups

The present NS-DA-SIM approach is now applied to PTV data obtained on a transitional round air jet [3]. The experimental apparatus is illustrated in figure 12a. Based on the exit diameter D and vertical velocity V at the center of the exit plane, which will be used to non-dimensionalize all quantities in this section, the corresponding Reynolds number is equal to $Re = 4600$. Measurements are obtained through the Double-Frame Tomographic PTV (DF-TPTV) approach which was proposed in [3]. In the following, we will consider the instantaneous data corresponding to figures 12 and 13 in [3], these measurements being here illustrated in figure 13a, which reports an extraction in a central slice. They correspond to relatively dense PTV measurements, i.e. to a seeding density of 0.06 particles per pixel, and a mean inter-particle distance $\overline{d} = 0.06$. These data are available in the domain that is depicted in figure 12b, which is also

used as computational domain in the data assimilation procedure. The origin O of the frame lies at center of the jet in the exit plane. The extents of the measurement domain in the horizontal and vertical (streamwise) directions are $L_x = 5$ and $L_y = 7.5$, respectively. In the z direction, the finite thickness of the illumination restricts the results to $-L_1 \leq z \leq L_2$, with $L_1 = 0.6$ and $L_2 = 0.4$, i.e. corresponding to an extent of one characteristic jet diameter but with a slight truncation of the jet core on one side.

The computational domain (which thus identities with the measurement one) is discretized with a mesh that is obtained by extruding the 2D mesh in figure 12c in the streamwise direction, resulting in a mesh of tetrahedra with a total of $7.1 \cdot 10^4$ vertices. Dirichlet boundary conditions $\boldsymbol{u} = \boldsymbol{0}$ are imposed at $x = \pm L_x/2 = 2.5$. This seems a good approximation given the values of the PTV data close to theses boundaries, and that the latter are already relatively far from the jet core. The other boundary conditions are determined from the first-guess velocity field (see §2.3), except at plane $y = L_y$ where the outflow condition $(\frac{\partial u}{\partial y}, Re^{-1}\frac{\partial v}{\partial y} - p, \frac{\partial w}{\partial y}) = \boldsymbol{0}$ is employed.

5.2. Reconstruction results

The NS-DA-SIM procedure is first performed relying on the full set of instantaneous data which are illustrated in figure 13a, the corresponding results being reported in figures 13-16. Note that we will also present results from a second reconstruction, based on only a partial set of these measurements. In the following descriptions, figures and tables, these two cases will be also referred to as 'full dataset' and 'partial dataset', respectively. The vertical velocity field v of the full dataset reconstruction is first illustrated in the $z = 0$ plane in figures 13b-13c for the first-guess flow and the reconstructed one, respectively. Despite the relatively modest decrease ($\sim 30\%$) in the value of the cost function J in this case compared to the previous synthetic cylinder flow (see figures 14 and 3b for comparison), it appears from figures 13b-13c that the data assimilation procedure has significantly modified the first-guess flow. In particular, it has sharpened velocity gradients at the boundaries of the jet core compared to the smoother first-guess flow. In order to confirm the quality of the reconstructed flow, the latter is compared in figures 13d-13e with mean profiles $\bar{v}$ of the vertical velocity at $y = 1.4$ and $y = 2$, which are obtained through both time and azimuthal averaging of DF-TPTV results using spatial bins in the form of annuli centered with the jet axis [3]. As the present reconstructed flow corresponds to a single instant, only azimuthal average is performed on the latter and on the first-guess flow for this comparison. Figures 13d-13e confirm that the first-guess flow significantly underestimates velocity gradients, in particular at a distance from the jet core $r = 0.5$, i.e. one characteristic jet radius. In contrast, the output of the NS-DA-SIM approach appears to well capture the highest gradient zones. It is actually much closer to the mean profiles $\bar{v}$ at all r, and remaining differences are seen to be of the same order of magnitude as the root-mean-square fluctuating velocity v_{rms}, which can be taken as an additional confirmation of the quality of the reconstructed flow.

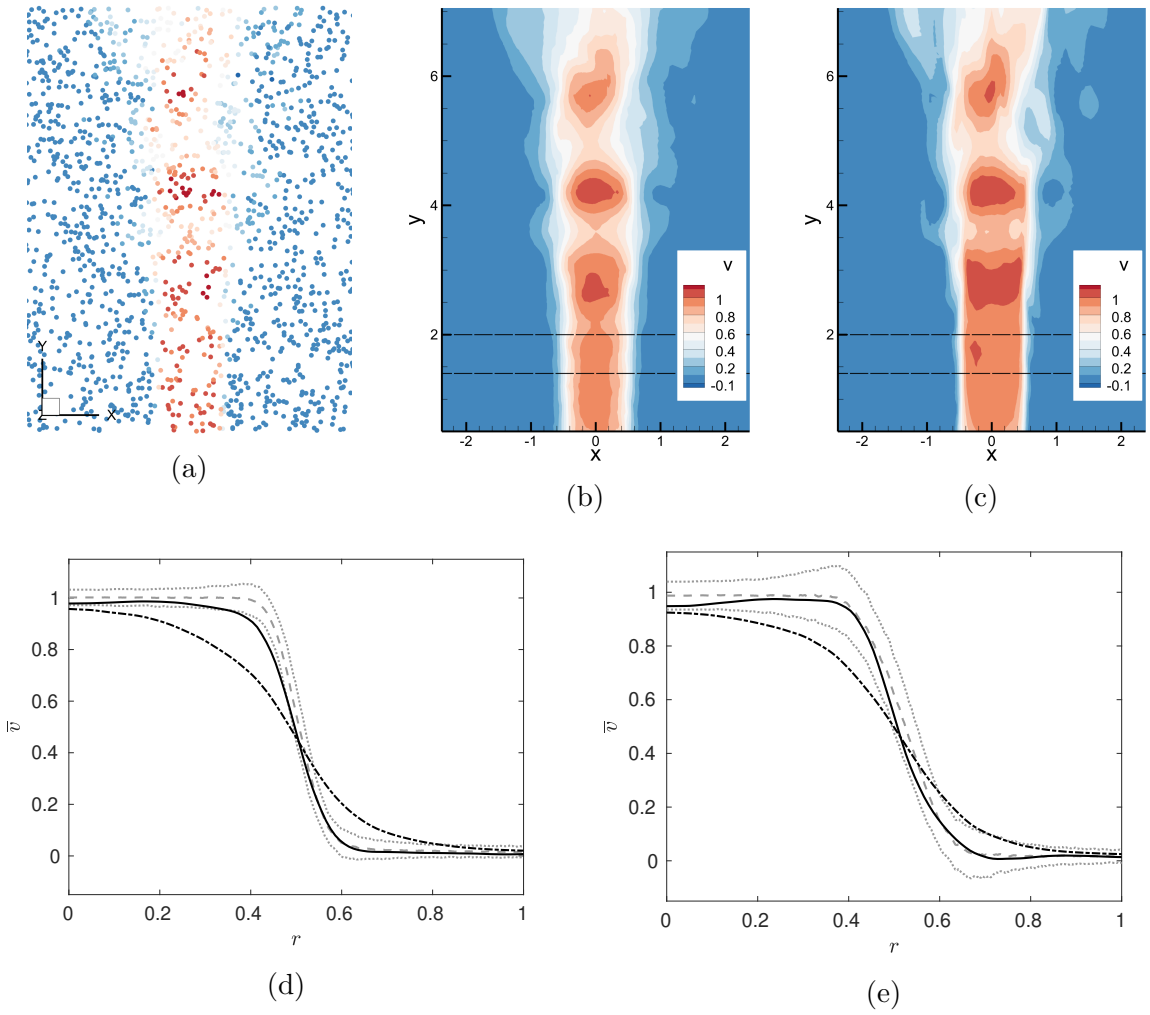


Figure 13: Data assimilation using the full PTV measurement set. (a) Measurement locations in z-normal view colored by the vertical velocity v in a $-0.025 \leq z \leq 0.025$ slice. (b)-(c) v field in the $z = 0$ plane for (b) the first-guess flow and (c) the reconstructed flow from the NS-DA-SIM approach. (d)-(e) Averaged $\bar{v}$ as a function of the distance r to the center of the round jet at (d) $y = 1.4$ and (e) $y = 2$, these y-locations being denoted by dashed lines in (b)-(c). Grey lines correspond to azimuthal and time bin averaging of PTV data, with profiles of $\bar{v}$ (- - - -) and $\bar{v} \pm v_{\text{rms}}$ (·····). Black lines correspond to the azimuthal average of the estimated instantaneous v field for the first-guess flow (-·-·-) and the reconstructed one (——).

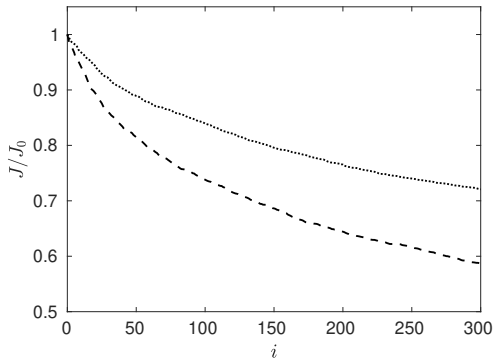


Figure 14: Cost function J (normalized by its initial value J_0) versus the iteration of the minimization procedure in the NS-DA-SIM approach when all (·····, full set) or only one third (- - - , partial set) of the jet flow PTV measurements are used, respectively.

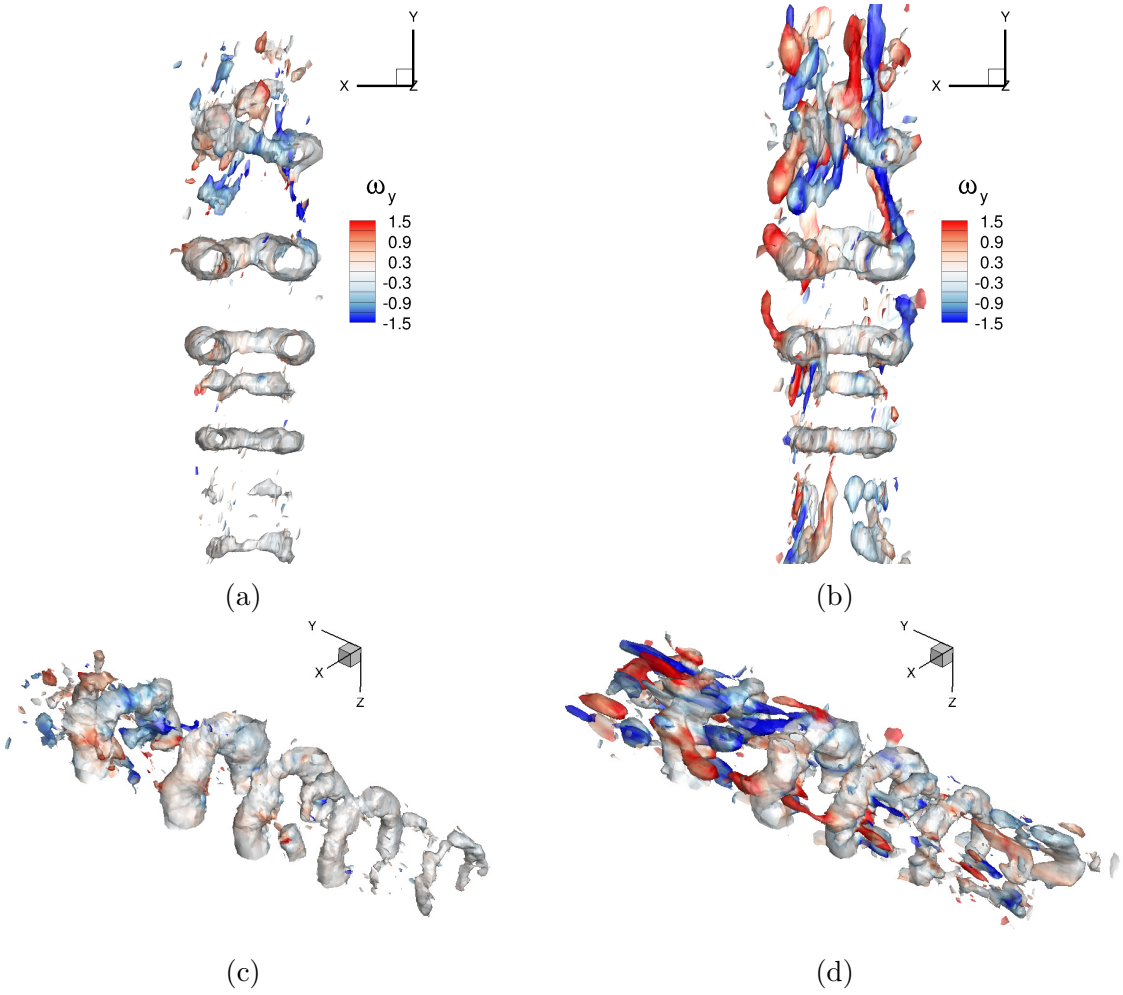
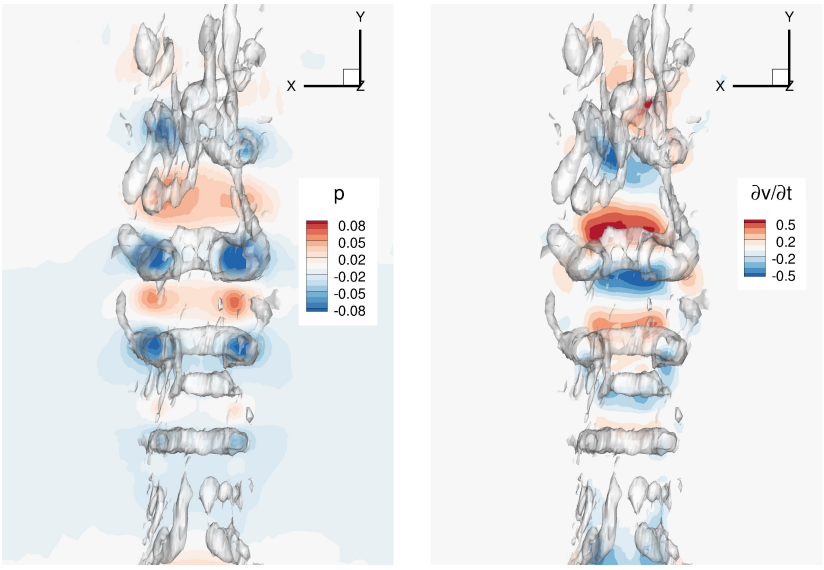


Figure 15: Q-criterion isosurface ($Q = 0.4$) colored by the streamwise vorticity ω_y for (a,c) the first-guess flow and (b,d) the reconstructed flow from the NS-DA-SIM approach (full dataset case) from two perspectives (top and bottom).



(a) (b)

Figure 16: Isosurface $Q = 0.4$ with the (a) pressure p and (b) streamwise Eulerian acceleration $\frac{\partial v}{\partial t}$ in the $z = 0$ plane for he reconstructed flow from the NS-DA-SIM approach (full dataset case).

The enhancement of the flow estimation through the data assimilation procedure is further assessed through figure 15 where an isosurface of the Q-criterion colored by the streamwise vorticity ω_y is reported for both first-guess and reconstructed flows. While the first-guess flow mostly exhibits vortex rings, with only irregular traces of secondary streamwise vortices in its downstream part, the reconstructed flow is characterized by a much richer content in vortical structures. The vortex rings appear in a more definite form, with, for some of them, clearer signs of azimuthal corrugations which are typically observed in the near-field of round jets. A much larger number of well-definite streamwise vortices of alternate signs can also be evidenced, not only in the most downstream part of the jet where they are characterized by a significant spatial extent, but also more upstream where they appear with finer sizes. Their location relative to the vortex rings is fully consistent physically, i.e. mostly in between two subsequent vortex rings, which defines the so-called 'braid region' (see, e.g., [48]). These observations clearly further confirm the ability of the NS-DA-SIM procedure in providing a wealth of detailed information about high-gradient flow regions. The pressure and streamwise Eulerian acceleration fields are also illustrated in figure 16 for the reconstructed flow. Regions of lower pressure well coincide with the core of the vortex rings, as was expected. The spatial field of the Eulerian acceleration $\frac{\partial v}{\partial t}$ also seems physically sound, in particular with respect to the vortex rings. These rings indeed induce a local constriction of the jet column, which results, due to mass conservation, in a local increase of streamwise velocity in its core. Conversely, the region between two subsequent rings is characterized by a broadening of the jet core, and thus smaller streamwise velocities. Now, recalling

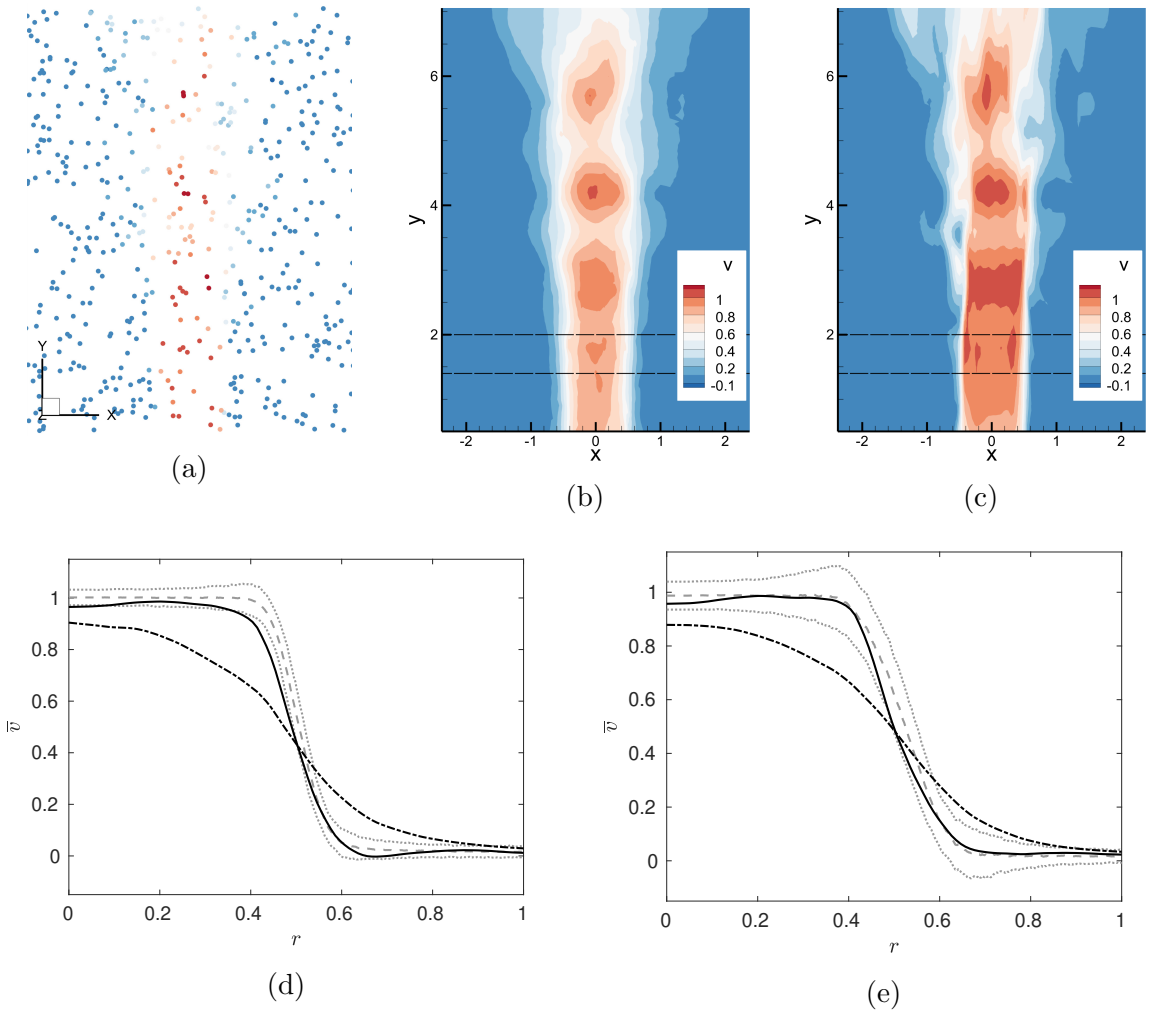
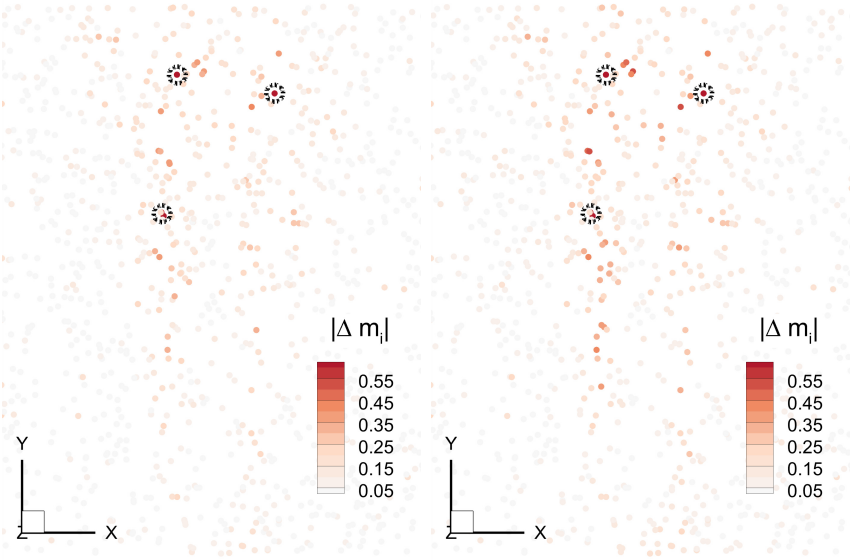


Figure 17: Data assimilation using one third of the jet flow PTV measurements (partial dataset case). The legend is otherwise the same as for figure 13.

that these rings have a motion of translation in the streamwise direction, one can understand that the presence of positive (resp. negative) patches of $\frac{\partial v}{\partial t}$ downstream (resp. upstream) of a given ring is consistent. Indeed, the region directly upstream of a ring must be undergoing a progressive constriction (thus acceleration), while the region downstream must be in a phase of broadening (thus deceleration).

The NS-DA-SIM approach is here further assessed in a second case where only one third of the available measurements is considered (partial dataset case). The corresponding results are reported in figure 17. The mean inter-particle distance becomes $\overline{d} = 0.09$ (instead of $\overline{d} = 0.06$ previously), corresponding to a flow seeding density of 0.02 particles per pixel. The decrease in the measurement density has a significant impact on the first-guess flow, which in particular appears to further underestimate the intensity of streamwise velocity in the acceleration regions, as may be inferred from the comparison between figures 13b and 17b. On the other hand, although some differences are visible, the reconstructed flow in the present case (see



(a) (b)

Figure 18: Discrepancies $|\Delta m_i| = \|\boldsymbol{u}(\boldsymbol{x}_i) - \boldsymbol{m}_i\|$ between PTV measurements and the reconstructed flow from the NS-DA-SIM approach in a $-0.025 \leq z \leq 0.025$ slice. The two thirds of the data that are not used in the partial dataset case are considered in this figure, for the (a) full and (b) partial dataset case, respectively. Circled particles correspond to $|\Delta m_i| > 0.6$ in both cases.

figure 17c) appears close to the one in the previous data assimilation procedure based on the full measurements (see figure 13c). The still very satisfactory comparison between the present reconstructed flow and the mean profiles in figures 17d-17e further supports the robustness of the steady NS-DA-SIM approach with respect to the measurement scarcity in this case.

As a supplementary validation of the NS-DA-SIM methodology, figure 18 reports the discrepancies $|\Delta m_i| = \|\boldsymbol{u}(\boldsymbol{x}_i) - \boldsymbol{m}_i\|$, where $\|\circ\|$ denotes the Euclidean norm on $\mathbb{R}^3$, between the reconstructed flow and the remaining two thirds of the full measurements that were not used in the partial dataset case. Note that we represent this quantity for the two reconstructions presented above, i.e. both when using the full and partial datasets in the assimilation procedure. The discrepancies presented in figure 18a for the full dataset case should correspond to the minimum residual discrepancies that may be reached at the reported locations, and will serve as a reference in the following. Values of $|\Delta m_i|$ are overall low, in particular outside of the jet and in the jet core. Errors seem larger in the shear layers and in the most downstream part of the jet, which correspond to regions with higher velocity gradients. Interestingly, discrepancies $|\Delta m_i|$ associated to particles adjacent to a particle characterized by a high $|\Delta m_i|$ are generally found significantly lower. In other words, high values of $|\Delta m_i|$ appear isolated in space, which may suggest that these particles could be residual outliers in the PTV dataset that were not filtered out by the statistical rejection used in [3]. This might be the

Table 1: Average discrepancies E (as defined by (19)) between PTV measurements and reconstructed flows obtained by performing the NS-DA-SIM procedure with full and partial measurement datasets, evaluated by considering either the full data (N_{PTV}), the partial dataset ($N_{1/3}$), or its complementary ($N_{2/3}$). This quantity is also evaluated by considering the same data ensembles, but discarding measurements corresponding to the 5% largest discrepancies of the full dataset case ($\widetilde{\cdot}$).

	N_{PTV}	$N_{1/3}$	$N_{2/3}$	$\widetilde{N_{\mathrm{PTV}}}$	$\widetilde{N_{1/3}}$	$\widetilde{N_{2/3}}$
full dataset	$1.0 \cdot 10^{-2}$	$1.0 \cdot 10^{-2}$	$1.0 \cdot 10^{-2}$	$6.3 \cdot 10^{-3}$	$6.3 \cdot 10^{-3}$	$6.3 \cdot 10^{-3}$
partial dataset	$1.3 \cdot 10^{-2}$	$1.0 \cdot 10^{-2}$	$1.5 \cdot 10^{-2}$	$8.3 \cdot 10^{-3}$	$6.5 \cdot 10^{-3}$	$9.3 \cdot 10^{-3}$

case in particular for exemplified (circled) particles in figure 18a, which correspond to the largest discrepancies, with $|\Delta m_i| > 0.6$. Now turning to the partial dataset case, which is illustrated in figure 18b, the comparison between this figure and figure 18a does not suggest a drastic increase in the errors with respect to the unused measurements, compared to the full dataset case. Interestingly, the largest discrepancies are reached at the same locations in both cases, which may be a supplementary indication of the presence of outliers there.

A quantitative assessment of these discrepancies is provided in table 1 where we report the average error

$$E = \frac{1}{N} \sum_{i=1}^{N} \|\boldsymbol{u}(\boldsymbol{x}_i) - \boldsymbol{m}_i\|^2, \tag{19}$$

where the sum is performed over the whole set of measurements (N_{PTV}), the partial dataset ($N_{1/3}$) or its complementary ($N_{2/3}$). Motivated by the above discussion on the possible presence of outliers, the error E is also evaluated when excluding the measurements that correspond to the 5% largest remaining discrepancies of the full dataset case. Concerning this first case, table 1 confirms the similar error level in the two measurement subsets (see columns $N_{1/3}$ and $N_{2/3}$). It also indicates that by removing only the 5% most problematic measurements, E is decreased from $1.0 \cdot 10^{-2}$ to $6.3 \cdot 10^{-3}$, thus by 37%. Concerning the partial dataset case, it reaches the same discrepancy levels at the used measurements (see column $N_{1/3}$) as in the full dataset case, while errors with respect to the unused measurements (see column $N_{2/3}$) are 50% larger. These discrepancies still remain low, which further supports the efficacy of the NS-DA-SIM approach.

6. Conclusions

In this study, a variational data assimilation approach for instantaneous 3D flow reconstruction has been proposed. It has been specifically designed to rely on single-instant velocity measurements only, and therefore does not require any time-resolved

or acceleration data. It is based on a strong enforcement of the Navier-Stokes equations in a steady form where a forcing in the momentum equation, which should account for Eulerian acceleration, is adjusted to minimize the discrepancies between the reconstructed flow and the velocity measurements. Focusing on the case of scattered measurements as in PTV, potential spurious effects linked to the introduction of pointwise data are avoided thanks to regularization of the gradient of the cost function in the data assimilation procedure, resulting in a parameter-free methodology, since the length scale in this regularization can be taken as the mean distance between neighboring measurements.

The ability of the proposed NS-DA-SIM approach to accurately reconstruct full instantaneous velocity, pressure and Eulerian acceleration fields, along with estimating aerodynamic forces, has been thoroughly assessed based on the 3D numerical flow past a cylinder at $Re = 300$. Comparisons with alternative data assimilation schemes have further emphasized the merits of the present approach. The ability of the latter to handle volumetric PTV measurements has also been demonstrated for a transitional jet flow experiment at $Re = 4600$, as evidenced in particular by the rich and physical vortical content of the reconstructed flows, as well as the physical soundness of the pressure and Eulerian acceleration fields.

Obvious extensions of the present study include the application of the present methodology to other flow configurations. The present optimization framework could also be extended to the use of robust norms [49] to better handle outliers, which are unavoidable in any experimental context. More generally, measurement uncertainties/errors could be taken into account when available in the data assimilation procedure, possibly in an iterative process where data assimilation itself could be used to update the estimation of these uncertainties. An even stronger coupling with the PTV data processing could also be considered, also within an iterative approach, whereby the reconstructed flows would yield successive displacement predictors in order to detect and match the largest possible number of particles. Finally, the present formalism should be versatile enough to enable the consideration of other types of velocity measurements (e.g. PIV), but also of pressure or temperature, in conjunction with other models than the incompressible Navier-Stokes equations.

Appendix A. Other variational data assimilation approaches

Appendix A.1. Stokes-based data assimilation

Instead of relying on the (steady) Navier-Stokes equations, one could consider the strong imposition of the following Stokes equations, similarly as performed in [47] in the context of optical flow (i.e. working directly on particle images)

$$\mathbf{S}\boldsymbol{q} = \mathbf{P}\boldsymbol{g}, \quad \mathbf{S} = \begin{pmatrix} -Re^{-1}\Delta\circ & \nabla\circ \\ -\nabla\cdot\circ & 0 \end{pmatrix}. \tag{A.1}$$

The forcing $\boldsymbol{g}$ in the above equation should account for the opposite of the material

derivative of the velocity field, namely $-\frac{d\boldsymbol{u}}{dt} = -\frac{\partial \boldsymbol{u}}{\partial t} - (\boldsymbol{u}\cdot\nabla)\boldsymbol{u}$. It may noted that this quantity is not divergence-free, which should preclude an unambiguous reconstruction of the pressure field (see §2.4). The data assimilation problem based on the Stokes equations (A.1) may be expressed as

$$\min_{\boldsymbol{g}} \left\{ J = \frac{1}{2}\|\boldsymbol{m} - \mathbf{H}\boldsymbol{u}\|^2_{\mathcal{M}} \right\}, \quad \mathbf{S}\boldsymbol{q} = \mathbf{P}\boldsymbol{g}, \tag{A.2}$$

Similarly as in the NS-DA-SIM procedure, (A.2) is here solved through an iterative gradient-based algorithm. The gradient $\frac{dJ}{d\boldsymbol{g}}$ is obtained in a similar way as described in §2.2, the adjoint problem for the Stokes equations being close to (9)-(10) but dropping the first two contributions in the left-hand-side of the adjoint momentum equation which originate from the convection term. The identification of a first guess for this procedure is performed in a similar way as described in §3.1, only changing its third step and evaluating the left-hand-side of the momentum equation in (A.1) to get a first guess for $\boldsymbol{g}$.

Appendix A.2. Unsteady Navier-Stokes-based data assimilation (4DVar)

Variational data assimilation based on the unsteady Navier-Stokes equations (2), which was almost exclusively considered for time series of measurements (see, e.g., [20, 21, 24]), may be applied as follows in the present context of single-instant measurements [25]. The time dependency of the flow, which was discarded in §2.1 to simplify notations, is here reintroduced for the sake of clarity. Considering a time window of size T, which forms the main adjustable parameter of the following methodology, the data assimilation problem is formulated as identifying the initial condition $\boldsymbol{u}(t_m - T)$ so that the velocity field $\boldsymbol{u}(t_m)$ matches the measurements $\boldsymbol{m}(t_m)$. The corresponding optimization problem may be written as

$$\begin{aligned} &\min_{\boldsymbol{u}(t_m-T)} \left\{ J = \frac{1}{2}\|\boldsymbol{m}(t_m) - \mathbf{H}\boldsymbol{u}(t_m)\|^2_{\mathcal{M}} \right\}, \\ &\mathbf{P}\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{\mathcal{N}}(\boldsymbol{q}) = \mathbf{0}, \quad t \in [t_m - T, t_m], \end{aligned} \tag{A.3}$$

where, using (5), the second equality refers to the imposition of the unsteady Navier-Stokes equations. Similarly as in §2.2 and Appendix A.1, (A.3) is solved in an iterative way. The gradient of the cost function J with respect to the initial condition $\boldsymbol{u}(t_m - T)$ is obtained by first integrating backward in time from $t = t_m$ to $t = t_m - T$ the adjoint unsteady Navier-Stokes equations

$$-\mathbf{P}\frac{\partial \boldsymbol{u}^\dagger}{\partial t} + \mathbf{N}^\dagger \boldsymbol{q}^\dagger = \mathbf{0}, \quad t \in [t_m - T, t_m], \tag{A.4}$$

with the terminal condition $\boldsymbol{u}^\dagger(t_m) = \mathbf{H}^\dagger(\mathbf{H}\boldsymbol{u}(t_m) - \boldsymbol{m}(t_m))$. The definition of the involved adjoint operators may be found in §2.2. The required gradient is then obtained through $\frac{dJ}{d\boldsymbol{u}(t_m-T)} = \boldsymbol{u}^\dagger(t_m - T)$. It may be emphasized that the backward integration of the unsteady adjoint equations (A.4) first requires the integration of the unsteady Navier-Stokes and the storage (or the recomputation) of the full flow history in the

interval $[t_m - T, t_m]$. The identification of a first guess for the initial condition $\boldsymbol{u}(t_m - T)$ is performed by first following the first step described in §3.1, which provides a first guess for $\boldsymbol{u}(t_m)$. The unsteady Navier-Stokes equations are then solved backward in time, negating the diffusion term to ensure stable computations, to get the first-guess initial condition.

References

[1] G. E. Elsinga, F. Scarano, B. Wieneke, and B. W. van Oudheusden. Tomographic particle image velocimetry. Experiments in Fluids, 41:933–947, 2006.

[2] T. Fuchs, R. Hain, and C. J. Kähler. Double-frame 3d-ptv using a tomographic predictor. Experiments in Fluids, 57:1–5, 2016.

[3] P. Cornic, B. Leclaire, F. Champagnat, G. Le Besnerais, A. Cheminet, C. Illoul, and G. Losfeld. Double-frame tomographic PTV at high seeding densities. Experiments in Fluids, 61:23, 2020.

[4] B. W. van Oudheusden. PIV-based pressure measurement. Measurement Science and Technology, 24:032001, 2013.

[5] H. Gunes, S. Sirisup, and G. E. Karniadakis. Gappy data: To Krig or not to Krig? Journal of Computational Physics, 212:358–382, 2006.

[6] M. Raiola, S. Discetti, and A. Ianiro. On PIV random error minimization with optimal POD-based low-order reconstruction. Experiments in Fluids, 56:75, 2015.

[7] K. Fukami, K. Fukagata, and K. Taira. Super-resolution reconstruction of turbulent flows with machine learning. Journal of Fluid Mechanics, 870:106–120, 2019.

[8] A. Güemes, C. S. Vila, and S. Discetti. GANs-based PIV resolution enhancement without the need of high-resolution input. 14th International Symposium on Particle Image Velocimetry, 2021.

[9] J. M. Lewis, S. Lakshmivarahan, and S. K. Dhall. Dynamic data assimilation: a least squares approach, volume 104 of Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2006.

[10] R. Vennell and R. Beatson. A divergence-free spatial interpolator for large sparse velocity data sets. Journal of Geophysical Research: Oceans, 114:C10024, 2009.

[11] I. Azijli and R. P. Dwight. Solenoidal filtering of volumetric velocity measurements using Gaussian process regression. Experiments in Fluids, 56:198, 2015.

[12] J. J. J. Gillissen, R. Bouffanais, and D. K. P. Yue. Data assimilation method to de-noise and de-filter particle image velocimetry data. Journal of Fluid Mechanics, 877:196–213, 2019.

[13] J. F. G. Schneiders, S. Pröbsting, R. P. Dwight, B. W. van Oudheusden, and F. Scarano. Pressure estimation from single-snapshot tomographic PIV in a turbulent boundary layer. Experiments in Fluids, 57:53, 2016.

[14] J. F. G. Schneiders, F. Avallone, S. Pröbsting, D. Ragni, and F. Scarano. Pressure spectra from single-snapshot tomographic PIV. Experiments in Fluids, 59:57, 2018.

[15] A. Sciacchitano, R. P. Dwight, and F. Scarano. Navier–Stokes simulations in gappy PIV data. Experiments in Fluids, 53:1421–1435, 2020.

[16] T. Suzuki and F. Yamamoto. Hierarchy of hybrid unsteady-flow simulations integrating time-resolved PTV with DNS and their data-assimilation capabilities. Fluid Dynamics Research, 47: 051407, 2015.

[17] S. Gesemann, F. Huhn, D. Schanz, and A. Schröder. From Noisy Particle Tracks to Velocity, Acceleration and Pressure Fields using B-splines and Penalties. 18th International Symposium on the Application of Laser and Imaging Techniques to Fluid Mechanics, 2016.

[18] F. Ehlers, A. Schröder, and S. Gesemann. Enforcing temporal consistency in physically constrained flow field reconstruction with FlowFit by use of virtual tracer particles. Measurement Science and Technology, 31:094013, 2020.

[19] J. F. G. Schneiders and F. Scarano. Dense velocity reconstruction from tomographic PTV with material derivatives. Experiments in Fluids, 57:139, 2016.

[20] A. Gronskis, D. Heitz, and E. Mémin. Inflow and initial conditions for direct numerical simulation based on adjoint data assimilation. Journal of Computational Physics, 242:480–497, 2013.

[21] R. Yegavian, B. Leclaire, F. Champagnat, and O. Marquet. Performance assessment of PIV super-resolution with adjoint-based data assimilation. 11th International Symposium on Particle Image Velocimetry, 2015.

[22] M. Lemke and J. Sesterhenn. Adjoint-based pressure determination from PIV data in compressible flows - Validation and assessment based on synthetic data. European Journal of Mechanics B/Fluids, 58:29–38, 2016.

[23] P. Chandramouli, E. Mémin, and D. Heitz. 4D large scale variational data assimilation of a turbulent flow with a dynamics error model. Journal of Computational Physics, 412:109446, 2020.

[24] F. Scarano, J. F. G. Schneiders, G. Gonzalez Saiz, and A. Sciacchitano. Dense velocity reconstruction with VIC-based time-segment assimilation. Experiments in Fluids, 63:96, 2022.

[25] L.-P. Saumier, B. Khouider, and M. Agueh. Effective filtering and interpolation of 2D discrete velocity fields with Navier–Stokes equations. Inverse Problems, 32:115006, 2016.

[26] D. P. G. Foures, N. Dovetta, D. Sipp, and P. J. Schmid. A data-assimilation method for Reynolds-averaged Navier-Stokes-driven mean flow reconstruction. Journal of Fluid Mechanics, 759:404–431, 2014.

[27] S. Symon, N. Dovetta, B. J. McKeon, D. Sipp, and P. J. Schmid. Data assimilation of mean velocity from 2D PIV measurements of flow over an idealized airfoil. Experiments in Fluids, 58: 61, 2017.

[28] B. Protas, T. R. Bewley, and G. Hagen. A computational framework for the regularization of adjoint analysis in multiscale PDE systems. Journal of Computational Physics, 195:49–89, 2004.

[29] G. Tissot, R. Billard, and G. Gabard. Optimal cavity shape design for acoustic liners using Helmholtz equation with visco-thermal losses. Journal of Computational Physics, 402:109048, 2020.

[30] M. Y. Ben Ali. Investigating Data-Model coupling using adjoint techniques for wind engineering. PhD thesis, Université de Rennes 1, 2021.

[31] J. E. V. Peter and R. P. Dwight. Numerical sensitivity analysis for aerodynamic optimization: A survey of approaches. Computers & Fluids, 39:373–391, 2010.

[32] V. Mons and O. Marquet. Linear and nonlinear sensor placement strategies for mean-flow reconstruction via data assimilation. Journal of Fluid Mechanics, 923:A1, 2021.

[33] J. Nocedal. Updating Quasi-Newton Matrices With Limited Storage. Mathematics of Computation, 35:773–782, 1980.

[34] L. Armijo. Minimization of functions having Lipschitz continuous first partial derivatives. Pacific Journal of Mathematics, 16:1–3, 1966.

[35] L. Franceschini, D. Sipp, and O. Marquet. Mean-flow Data Assimilation based on minimal correction of turbulence models: Application to turbulent high-Reynolds number backward-facing step. Physical Review Fluids, 5:094603, 2020.

[36] P. C. Hansen and D. P. O'Leary. The use of the L-curve in the regularization of discrete ill-posed problems. SIAM Journal on Scientific Computing, 14:1487–1503, 1993.

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

[38] D. N. Arnold, F. Brezzi, and M. Fortin. A stable finite element for the Stokes equations. Calcolo, 21:337–344, 1984.

[39] A. N. Brooks and T. J. R. Hughes. Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations. Computer Methods in Applied Mechanics and Engineering, 32:199–259, 1982.

[40] M. Olshanskii, G. Lube, T. Heister, and J. Löwe. Grad–div stabilization and subgrid pressure models for the incompressible Navier–Stokes equations. Computer Methods in Applied Mechanics and Engineering, 198:3975–3988, 2009.

[41] A. Crivellini and F. Bassi. An implicit matrix-free Discontinuous Galerkin solver for viscous and turbulent aerodynamic simulations. Computers & Fluids, 115:81–93, 2011.

[42] R. Codina. Pressure Stability in Fractional Step Finite Element Methods for Incompressible Flows. Journal of Computational Physics, 170:112–140, 2001.

[43] H. Jiang, L. Cheng, S. Draper, H. An, and F. Tong. Three-dimensional direct numerical simulation of wake transitions of a circular cylinder. Journal of Fluid Mechanics, 801:353–391, 2016.

[44] C. H. K. Williamson. Vortex dynamics in the cylinder wake. Annual Review of Fluid Mechanics, 28:477–539, 1996.

[45] D. Barkley and R. D. Henderson. Three-dimensional Floquet stability analysis of the wake of a circular cylinder. Journal of Fluid Mechanics, 322:215–241, 1996.

[46] P. Parnaudeau, J. Carlier, D. Heitz, and E. Lamballais. Experimental and numerical studies of the flow over a circular cylinder at Reynolds number 3900. Physics of Fluids, 20:085101, 2008.

[47] P. Ruhnau and C. Schnörr. Optical Stokes flow estimation: an imaging-based control approach. Experiments in Fluids, 42:61–78, 2007.

[48] D. Liepmann and M. Gharib. The role of streamwise vorticity in the near-field entrainment of round jets. Journal of Fluid Mechanics, 245:643–668, 1992.

[49] P. J. Huber. Robust regression: Asymptotics, Conjectures and Monte Carlo. The Annals of Statistics, 1:799–821, 1973.