CrossMark

RESEARCH ARTICLE

Data assimilation of mean velocity from 2D PIV measurements of flow over an idealized airfoil

Sean Symon[1] · Nicolas Dovetta[2] · Beverley J. McKeon[1] · Denis Sipp[3] ·
Peter J. Schmid[4]

Abstract Data assimilation can be used to combine experimental and numerical realizations of the same flow to produce hybrid flow fields. These have the advantages of less noise contamination and higher resolution while simultaneously reproducing the main physical features of the measured flow. This study investigates data assimilation of the mean flow around an idealized airfoil ($Re = 13,500$) obtained from time-averaged two-dimensional particle image velocimetry (PIV) data. The experimental data, which constitute a low-dimensional representation of the full flow field due to resolution and field-of-view limitations, are incorporated into a simulation governed by the two-dimensional, incompressible Reynolds-averaged Navier–Stokes (RANS) equations with an unknown momentum forcing. This forcing, which corresponds to the divergence of the Reynolds stress tensor, is calculated from a direct-adjoint optimization procedure to match the experimental and numerical mean velocity fields. The simulation is projected onto the low-dimensional subspace of the experiment to calculate the discrepancy and a smoothing procedure is used to recover adjoint solutions on the higher dimensional subspace of the simulation. The study quantifies how well data assimilation can reconstruct the mean flow and the minimum experimental measurements needed

by altering the resolution and domain size of the time-averaged PIV.

1 Introduction

It has been noted in several studies (Nisugi et al. 2004; Suzuki et al. 2009a, b; Suzuki 2012; Foures et al. 2014) that despite recent advances in computational fluid dynamics (CFD) and experiments, both techniques have several disadvantages. Despite capturing the "true" physics of the flow, for example, experiments are corrupted by noise, limited by field of view, and have insufficient resolution to capture small scales. CFD, on the other hand, requires modeling assumptions about boundary conditions and sub-grid scale models unless there is sufficient computational power to resolve all scales in turbulence. Data assimilation seeks to combine the advantages of experiments and CFD to produce numerical simulations which obey the governing equations yet closely resemble the experimental data. The assimilated or hybrid flow is able to recover quantities in the experiment which would otherwise be inaccessible or difficult to measure such as pressure, vorticity, and Reynolds stresses by reducing noise and improving resolution. It is also possible to extrapolate the flow beyond the experimental view by solving the equations on a larger domain.

One of the first hybrid simulations conducted by Nisugi et al. (2004) used offline, sequential assimilation for flow behind a square cylinder. By measuring the discrepancy between experimental and numerical pressure measurements at finite time intervals to drive the momentum equations, the simulation was altered to match the experiment. Sequential assimilation was greatly expanded by Suzuki et al. (2009a, b) when particle-tracking velocimetry (PTV) data of an airfoil at high angle of attack was fed into a

✉ Sean Symon
ssymon@caltech.edu

1 GALCIT, California Institute of Technology, Pasadena, CA 91125, USA

2 LadHyX, Ecole Polytechnique, 91128 Palaiseau, France

3 ONERA-DAAA, 8 rue des Vertugadins, 92190 Meudon, France

4 Department of Mathematics, Imperial College London, London SW7 2AZ, UK

two-dimensional direct numerical simulation (DNS). The resulting hybrid flows contained less noise and recovered the unsteady pressure fields. They also offered insight into the statistics of the mean flow and the three-dimensional instabilities which attenuate vorticity strength.

Variational methods, which involve minimizing the distance between experimental and numerical velocity fields subject to governing equations, are another approach to data assimilation. The roots of variational data assimilation can be traced back to optimal control theory which has been applied to various flow control problems (see Kim and Bewley 2007, for an overview). Bewley et al. (2001), for example, studied the control side of the problem by investigating various control strategies applied to turbulent channel flow simulated using a DNS. Data assimilation, on the other hand, functions more closely to an estimator which reads in inputs from various sensors and fits them to an underlying model. The idea is to read in a sparse number of measurements and use the model to produce an estimated state which is more highly resolved in space and time. The current study investigates mean flow data assimilation, so improved time resolution is not applicable.

An improved estimation technique for mean flows has potential applications in mean flow modification studies. A large body of work has attempted to investigate this problem which has its roots in the experiments of Strykowski and Sreenivasan (1990), who showed experimentally that for low Reynolds numbers, a small control cylinder inserted in the wake behind a larger cylinder can completely suppress vortex shedding. Numerical studies including Giannetti and Luchini (2007) and Marquet et al. (2008) looked at the sensitivity of the cylinder instability to base flow modification and steady forcing near the critical Reynolds number of 47. Meliga et al. (2012) and Mettot and Sipp (2014) expanded this framework to higher Reynolds numbers and determined how a small control cylinder could impact the frequency of vortex shedding as predicted by the most unstable global mode of the mean flow. Data assimilation could expand control techniques to wall actuators such as an oscillating ribbon or synthetic jet which are difficult to model computationally due to ambiguous boundary conditions at the wall. It is possible, for example, to determine the mean flow from an experiment and recover a more highly resolved mean flow by tuning the boundary condition at the wall so that the simulated mean flow matches the experimental one.

A study conducted by Gronskis et al. (2013) illustrates this point quite well. They employed adjoint data assimilation to generate initial and inflow conditions for a DNS of flow around a cylinder at a Reynolds number of $Re = 172$. The resulting simulation reflected the flow physics from large-scale PIV measurements but contained far lower noise levels. Data assimilation has also been demonstrated
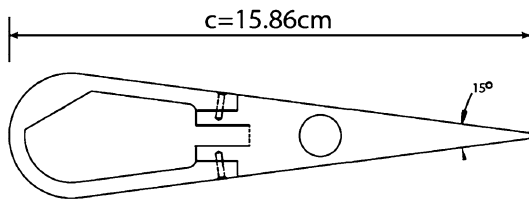
by Mons et al. (2016) to be applicable for perturbed fluid problems. They compared variational, ensemble Kalman filter-based, and ensemble-based variational data assimilation techniques to reconstruct the flow around a cylinder subject to coherent gusts. They found that the variational data assimilation approach produced the best results since the adjoint method can effectively capture the first-order sensitivity of the cost functional which penalizes the distance between experimental and numerical velocity fields.

The current study also employs variational data assimilation modeled on the approach taken by Foures et al. (2014), who applied variational methods to minimize the discrepancy between the mean velocity fields of a DNS and an incompressible RANS simulation for flow around a circular cylinder at a Reynolds number of $Re = 150$. The aim is to adapt this framework to mean flows obtained from experimental data at significantly higher Reynolds numbers since variational data assimilation has already been demonstrated in other forms for non-experimental data (e.g., synthetic PIV) or low Reynolds number flows. Experimental data are obtained from time-averaged particle image velocimetry (PIV) data from a free-surface water tunnel and integrated with a numerical model predicated on the two-dimensional incompressible Reynolds-averaged Navier–Stokes (RANS) equations with an unknown momentum forcing. This study investigates the mean flow at a chord-based Reynolds number of $Re = 13,500$ around an idealized airfoil consisting of a cylindrical leading edge followed by two plane surfaces connected at the trailing edge.

The outline of this paper is as follows: Sect. 2 describes the experimental setup and procedure used to obtain the data set. Next, Sect. 3 outlines the governing equations and data assimilation framework. The practical implementation and numerical methods behind the algorithm are explained in Sect. 4. The results of the data assimilation procedure and the minimum number of experimental measurements are explored in Sect. 5 while the residual discrepancy is discussed in Sect. 6. Finally, a method for incorporating measurement uncertainty into the assimilation procedure is suggested in Sect. 7 before the presentation of conclusions in Sect. 8.

2 Experimental setup

Experiments are performed on an idealized airfoil (see Fig. 1) with a chord-length of 15.86 cm, a width of 3.43 cm and a spanwise extent of 50.8 cm. The airfoil is symmetric about the chord and consists of a cylindrical leading edge followed by two plane surfaces connected at the trailing edge, which has a thickness of 0.15 cm. The diameter of the cylinder is equal to the width of the airfoil so that the junction between the cylinder and planar surfaces is half a
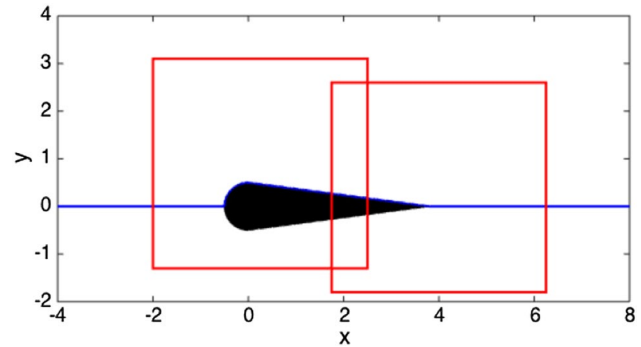
Fig. 1 Cross-sectional sketch of the idealized airfoil profile (Gonzalez et al. 2010)

Fig. 2 Experimental setup of the flow around an idealized airfoil showing the spatial coverage of the flow by the cameras. The dimensions are normalized by the diameter of the cylinder

diameter aft of the leading edge. At a zero angle of attack and a chord-based Reynolds number greater than approximately $Re = 10,000$, the flow separates at the transition point between the cylinder and the plane surfaces and the time-averaged flow reveals recirculation bubbles that form on both sides of the airfoil (Wallace and McKeon 2012). The critical Reynolds number, which coincides with the onset of von-Karman vortex shedding, is measured experimentally to be approximately 2000. Furthermore, the shear layers which form around the separation bubbles are convectively unstable, giving rise to the formation of Kelvin–Helmholtz vortices. These two instabilities ensure dynamics which are reminiscent of the behavior of separated flows at far higher Reynolds numbers (Prasad and Williamson 1997). Consequently, this configuration is an attractive choice for studying the capability of data assimilation to capture the flow dynamics around aerodynamic geometries.

Experiments are conducted in a free-surface water facility (see Wallace and McKeon 2012). The test section measures 1.6 m in length, 0.46 m in width and 0.5 m in height and the airfoil is mounted vertically so that its span is parallel to the test section height. The flow is conditioned by a perforated plate, a honey-comb mesh, three turbulence-reducing screens and a 4-to-1 fifth-order-polynomial contraction (Gharib 1983). The free-stream velocity is 8.1 cm/s and the free-stream turbulence intensity is less than 0.1% at the centerline. The water temperature is 23 °C, which results in a chord-based Reynolds number of $Re = 13,500$.

A LaVision time-resolved 2D-PIV setup is used consisting of two Photron Fastcam APX-RS high-speed cameras with 50 mm focal length Nikon lenses and 1:1.2 aperture. The cameras are synchronized with a high-speed controller and sample the flow at a frame rate of 83 Hz. The camera resolution is 1024×1024 pixels and the cameras are calibrated at 5.25 px/mm. The snapshot frequency is selected to guarantee a particle displacement between 5 and 7 pixels between any two consecutive snapshots. The seeding particles are hollow glass spheres (reference 110P8 with an average diameter of 11.7 µm and a specific gravity of 1.1) and the seeding density is about 0.1 particles per square pixel. The particles are illuminated by a 2 mm-thick laser sheet provided by a Photonics DM20-527 solid-state

laser. In an effort to avoid large uncertainty near the illuminated profile due to surface reflection, the image intensity is calibrated using white-image subtraction (normalization of the image intensity using the average light distribution) and background-image subtraction. The white and background images are taken before each run and averaged over 100 snapshots. The camera view, shown in Fig. 2, encompasses the flow from 6.5 cm upstream of the leading edge to 10.5 cm downstream of the trailing edge with a 15% overlap to include a large area of mean flow and fluctuation measurements. Finally, 10,240 instantaneous flow fields are captured over 5 runs (2048 snapshots per run), which represents approximately 35 complete vortex shedding cycles.

The computation of velocity vectors is performed using the software package DaVis provided by LaVision. A standard cross-correlation technique via Fast Fourier Transformation (see Adrian 1991) is applied to each sequential image with a window size reduced from 32×32 px^2 to 16×16 px^2 over three passes, a 50% overlap, and a 2:1 elliptic weight (see Kompenhans et al. 2007). Finally, the data are post-processed and single missing vectors are interpolated using an average of all the non-zero neighborhood vectors. A median filter, as described by Westerweel and Scarano (2005), is used for outlier detection. This results in a total of 8424 two-dimensional velocity vectors per instantaneous flow field and a resolution of 2.65 mm/vector. The snapshots are then averaged to obtain the mean velocity field and Reynolds stresses.

The mean velocity field containing all the measurement vectors is used as a reference to evaluate the quality of the assimilated velocity fields. To investigate the minimum experimental data necessary for successful mean flow reconstruction, the experimental data are artificially limited to include fewer measurement points. Truncating the field of view or omitting vectors from the full data set can be accomplished without repeating the experiment or data post-processing as explained in Sect. 4.

3 Data assimilation framework

The primary elements of the data assimilation framework are similar to those outlined in Foures et al. (2014), and so only an abridged version is presented here. There are, nevertheless, several modifications that shall be discussed in the following section to account for a mean velocity data set that comes from experiments instead of a time-averaged DNS and a Reynolds number that is two orders of magnitude higher than the Reynolds number used in Foures et al. (2014). This section outlines the governing equations which serve as a computational model for the flow, a fitting criterion which quantifies the progress of the data assimilation procedure, and the adjoint equations which reveal gradient information that is subsequently incorporated into the optimization procedure.

3.1 Computational model

The measured mean flow is obtained from 2D PIV data and is assumed to satisfy the 2D incompressible RANS equations given by

$$\overline{\mathbf{u}} \cdot \nabla \overline{\mathbf{u}} + \nabla \overline{p} - \frac{1}{Re} \nabla^2 \overline{\mathbf{u}} = \mathbf{f} \tag{1a}$$

$$\nabla \cdot \overline{\mathbf{u}} = 0. \tag{1b}$$

The forcing term $\mathbf{f}$ on the right-hand side of Eq. (1a) is an unknown momentum forcing which is data-driven whereas for ordinary RANS simulations it is solved for using a turbulence model. It corresponds to the divergence of the Reynolds stress tensor, defined as

$$\mathbf{f} = -\nabla \cdot \mathbf{R}, \quad \text{with } R_{ij} = \overline{u_i' u_j'}, \tag{2}$$

where u' and v' denote the streamwise and transverse fluctuating velocity components, respectively. Out-of-plane velocities are not captured by the experiment which means enforcing Eq. (1b) may not strictly be true at every point in the domain. Non-zero three-dimensional mean flow effects are compensated for by the momentum forcing term (for a better treatment of 3D effects, see conclusion).

This algorithm is formulated on velocity-only measurements, so only a partial recovery of the pressure is possible by data assimilation of the mean velocity field. Upon further inspection of Eq. (1a), it is noted that the forcing term $\mathbf{f}$ can be decomposed into the following

$$\mathbf{f} = \underbrace{\nabla \phi}_{\mathbf{f}_i} + \underbrace{\nabla \times \psi}_{\mathbf{f}_s} \tag{3}$$

where $\mathbf{f}_i$ and $\mathbf{f}_s$ represent the irrotational part and solenoidal part of $\mathbf{f}$, respectively. As explained in greater detail by Foures et al. (2014), the recovered forcing is divergence-free; therefore, only the solenoidal part can be captured. Furthermore, boundary conditions need to be specified for this decomposition to be unique. This involves enforcing $\nabla \phi \cdot \mathbf{n} = 0$ and $\mathbf{f}_s \cdot \mathbf{n} = 0$ on the airfoil walls where $\mathbf{n}$ is the outward normal. The model given by Eq. (1) can now be recast as follows:

$$\overline{\mathbf{u}} \cdot \nabla \overline{\mathbf{u}} + \nabla \overline{p}' - \frac{1}{Re} \nabla^2 \overline{\mathbf{u}} = \mathbf{f}_s \tag{4a}$$

$$\nabla \cdot \overline{\mathbf{u}} = 0, \tag{4b}$$

where $\overline{p}' = \overline{p} - \phi$. Despite the limitation that the pressure field cannot be reconstructed in its entirety, it is still possible to reconstruct the mean velocity field using experimental data to correctly model $\mathbf{f}_s$. The portion of $\mathbf{f}$ which is not captured is lumped in with the pressure term.

3.2 Fitting criterion

The goal of the data assimilation algorithm is to determine the solenoidal forcing $\mathbf{f}_s$ which best matches the mean velocity fields from the simulation and experiment. To determine this forcing, it is first necessary to calculate the discrepancy between the measured flow $\overline{\mathbf{u}}_{\text{exp}}$ and the simulated flow $\overline{\mathbf{u}}$. The discrepancy velocity field is computed using

$$\Delta \mathbf{u} = \overline{\mathbf{u}}_{\text{exp}} - \overline{\mathbf{u}}. \tag{5}$$

Note that $\Delta \mathbf{u}$ is a two-dimensional vector containing both streamwise and transverse discrepancy velocity measurements for each mesh point. This computation is not trivial since the experimental and numerical data are defined on different points. A detailed explanation for how to compute this field is given in the next section. For now, it is assumed that the discrepancy field has the same spatial resolution as the numerical field and that it is possible to perform the above operation using techniques which do not involve interpolation.

The L_2-norm of the discrepancy field yields a scalar function called the fitting criterion, which quantifies the distance between the numerical simulation and the measured mean velocity. It is calculated using the following operation:

$$\mathcal{E}(\overline{\mathbf{u}}) = \frac{1}{2} \sum_{j=1}^{N} \left\| \Delta \mathbf{u}_j \right\|^2, \tag{6}$$

where N denotes the number of points on the mesh and the index j denotes the jth point on the mesh.

3.3 Cost functional and adjoint equations

Similar to 3D-Var (see Lewis et al. 2006), an objective functional using a variational formulation yields an iterative optimization scheme that minimizes the fitting criterion. The final result is an optimal match of the model-predicted data $\overline{\mathbf{u}}$ and the measured data $\overline{\mathbf{u}}_{exp}$. An augmented Lagrangian is formed consisting of the objective functional (the fitting criterion) and the constraints (the model equations) which are enforced in a weak form by Lagrange multipliers or adjoint variables. For readers familiar with optimal control, the augmented Lagrangian can also be thought of as a least-squares cost functional which penalizes the difference between the computed state, which in this case is $\overline{\mathbf{u}}$, and the observed state, $\overline{\mathbf{u}}_{exp}$. The cost functional is written as follows:

$$\mathcal{L}(\mathbf{f}_s, \overline{\mathbf{u}}, \mathbf{u}^\dagger, \overline{p}', p^\dagger) = \mathcal{E}(\overline{\mathbf{u}}) - \left\langle \mathbf{u}^\dagger, \overline{\mathbf{u}} \cdot \nabla \overline{\mathbf{u}} + \nabla \overline{p}' \right.$$
$$\left. - \frac{1}{Re} \nabla^2 \overline{\mathbf{u}} - \mathbf{f}_s \right\rangle - \left\langle p^\dagger, \nabla \cdot \overline{\mathbf{u}} \right\rangle. \quad (7)$$

The scalar product denoted by $\langle ., . \rangle$ is associated with the Euclidian norm for vector and scalar fields on the domain Ω.

To minimize the functional, first-order variational derivatives are taken with respect to each independent variable and set equal to zero. Differentiation with respect to the adjoint variables $\mathbf{u}^\dagger$ and $p^\dagger$ yields the direct equations given by Eq. (4) while differentiation with respect to the direct variables produces the adjoint equations given below:

$$-\overline{\mathbf{u}} \cdot \nabla \mathbf{u}^\dagger + \mathbf{u}^\dagger \cdot \nabla \overline{\mathbf{u}}^T - \nabla p^\dagger - \frac{1}{Re} \nabla^2 \mathbf{u}^\dagger = \frac{\delta \mathcal{E}}{\delta \overline{\mathbf{u}}} \quad (8a)$$

$$\nabla \cdot \mathbf{u}^\dagger = 0. \quad (8b)$$

The right-hand term of Eq. (8a) is discussed in the following section as it pertains to the smoothing procedure needed to recover adjoint solutions on the higher dimensional subspace of the simulation.

Finally, the variational derivative with respect to the forcing $\mathbf{f}_s$ yields

$$\nabla_{\mathbf{f}_s} \mathcal{E} = \mathbf{u}^\dagger, \quad (9)$$

which is the steepest descent direction towards the optimality condition. It can also be interpreted as the sensitivity of the fitting criterion to the forcing vector. An initial guess is necessary to begin the optimization procedure and compute the first descent direction. Since the forcing must be divergence-free and the Reynolds stresses must vanish on the airfoil surface, a natural initial guess is $\mathbf{f}_s = 0$. The solution to Eq. (8) then provides $\mathbf{u}^\dagger$, which is the direction in which

the guess to the forcing is updated. This can be stated mathematically by the following equation:

$$\mathbf{f}_{n+1} = \mathbf{f}_n + \alpha_n \mathbf{u}_n^\dagger, \quad (10)$$

where $\mathbf{f}_{n=0}$ denotes the initial guess. This method is known as the simple gradient descent method which, while effective, has shown to be prone to zigzagging since the new search direction is always orthogonal to the previous search direction (see Bewley et al. 2001). To avoid this problem, a conjugate-gradient approach using the Polak–Ribière formula (see Polak and Ribière 1969) is adopted as well as a line-search algorithm to determine α_n for each new descent direction.

As seen from Eq. (10), the forcing is a linear combination of adjoint velocity fields which, from Eq. (8b), must be divergence-free. This reinforces the validity of the initial guess and justifies the earlier statement that only the solenoidal component of the forcing can be captured by the data assimilation algorithm.

4 Practical implementation of the algorithm

This section explains the numerical and computational details for solving the equations outlined in Sect. 3. It is here that the modifications to the procedure by Foures et al. (2014) are introduced.

4.1 Numerical methods

Both the numerical and experimental data have been non-dimensionalized to yield a unit input velocity and a unit cylindrical diameter of the idealized airfoil. The direct and adjoint equations are solved using the finite-element software FreeFem++ (see http://www.freefem.org) to generate a two-dimensional triangular mesh. The computational domain Ω spans $-20 \leq x \leq 35$, $-12.5 \leq y \leq 12.5$ with the cylinder at the leading edge of the airfoil being centered at the origin as it is in the PIV domain shown in Fig. 2. The mesh density, which is controlled by specifying the number of divisions along a boundary, increases from the outer boundaries towards the airfoil surface. The equations are spatially discretized using quadratic basis functions for the velocity and linear basis functions for the pressure resulting in approximately 580,000 degrees of freedom for both velocity and pressure.

FreeFem++ solves the direct and adjoint equations in weak form which can be obtained by dotting the equations with a test function and integrating by parts to remove second derivative terms. The direct equations are solved using a Newton–Raphson method while the adjoint equations are linear in $\mathbf{u}^\dagger$ and can be solved efficiently without an iterative method. The boundary conditions for Eq. (4) are

no-slip on the airfoil surface, uniform velocity at the inlet, symmetry condition $\partial_y u = 0$, $v = 0$ on the upper and lower boundaries, and outflow condition $-\bar{p}'\mathbf{n} + Re^{-1}\nabla\bar{\mathbf{u}} \cdot \mathbf{n} = 0$ where $\mathbf{n}$ is the normal unit vector pointing out of the domain. The adjoint boundary conditions, which are determined via integration by parts are $\mathbf{u}^\dagger = 0$ at the inlet and on the airfoil surface, symmetry condition $\partial_y u^\dagger = 0$, $v^\dagger = 0$, and $Re^{-1}\partial_x u^\dagger + p^\dagger = -\bar{u}u^\dagger$, $Re^{-1}\partial_x v^\dagger = -\bar{u}v^\dagger$ at the outlet.

The boundary conditions, along with the governing equations, serve as constraints for the optimization procedure and are not a free parameter. The reader is referred to Gronskis et al. (2013) for an example of data assimilation where the inflow boundary condition for cylinder flow is a tunable parameter of the underlying model equations.

4.2 Smoothing procedure and choosing experimental points

It was assumed in the previous section that the resolutions of the experimental and numerical velocity fields are identical when computing the discrepancy velocity field. This is not the case since the experimental velocity field is defined on a coarse, uniform Cartesian grid while the numerical velocity field is continuous and approximated with quadratic basis functions. To obtain a discrete representation of the numerical velocity field, the values of the velocity defined on the vertices and midpoints of the finite-element triangles can be outputted from the finite-element code. A smoothing procedure is used to cope with the difference in resolution when forcing the adjoint RANS equations. For convenience, the points where experimental measurements exist shall be referred to as grid points while the points where numerical measurements are taken shall be referred to as mesh points. The smoothing procedure begins by dividing the domain into cells as seen in Fig. 3, where the red dot denotes the center of each cell and the blue vector the associated PIV measurement. Mesh points are sorted into the cells whose boundaries are delineated by the grey lines in Fig. 3. This means that for a given cell,

there is a unique experimental measurement and multiple numerical measurements. The fitting criterion in Eq. (6) can now be recast in terms of a new scalar product defined in the experimental measurement space:

$$\mathcal{E}(\bar{\mathbf{u}}) = \frac{1}{2}\langle \Delta\mathbf{u}, \Delta\mathbf{u}\rangle_M, \tag{11}$$

where the subscript M represents the number of measures in the inner product defined by

$$\langle \Delta\mathbf{u}_1, \Delta\mathbf{u}_2\rangle_M = \sum_{j=1}^{M} \Delta\mathbf{u}_{1,j}\Delta\mathbf{u}_{2,j}. \tag{12}$$

The numerical measurements are projected down to the measurement space using a smoothing function or operator $\mathcal{P}$:

$$\Delta\mathbf{u} = \bar{\mathbf{u}}_{\text{exp}} - \mathcal{P}\bar{\mathbf{u}}. \tag{13}$$

For this configuration, $\mathcal{P}$ was chosen to be an area average of the numerical measurements in a given cell. For the jth cell, therefore, $\mathcal{P}$ can be written as

$$\mathcal{P}_j\bar{\mathbf{u}} = \int_\Omega \bar{\mathbf{u}}(x)\phi_j(x)d\Omega, \tag{14}$$
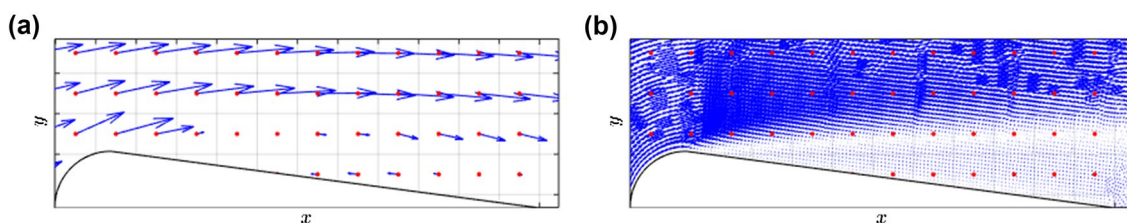
where

$$\phi_j(x) = \begin{cases} \frac{1}{\int_{\Omega_j} d\Omega} & \text{in } \Omega_j \\ 0 & \text{outside of } \Omega_j, \end{cases} \tag{15}$$

and $\int_{\Omega_j} d\Omega$ is the area of a cell.

The forcing term of Eq. (8a) can be derived by substituting Eq. (13) into Eq. (11) and taking the variational derivative with respect to $\bar{\mathbf{u}}$. After simplifying the expansion, one obtains

$$\left\langle \frac{\delta\mathcal{E}}{\delta\bar{\mathbf{u}}}, \delta\bar{\mathbf{u}}\right\rangle = \langle \bar{\mathbf{u}}_{\text{exp}} - \mathcal{P}\bar{\mathbf{u}}, -\mathcal{P}\delta\bar{\mathbf{u}}\rangle_M = \langle \mathcal{P}^\dagger(\mathcal{P}\bar{\mathbf{u}} - \bar{\mathbf{u}}_{\text{exp}}), \delta\bar{\mathbf{u}}\rangle, \tag{16}$$

and so $\frac{\delta\mathcal{E}}{\delta\bar{\mathbf{u}}}$, therefore, reads

Fig. 3 Subdivision of the domain into cells delineated by *grey lines*. The center of each cell is indicated by a *red point* which corresponds to the spatial location of a PIV measurement shown by a *blue vector*. Only a limited section of the total field of view is shown. In (**a**) one out of every five PIV vectors in both the streamwise and transverse directions appear for clarity while (**b**) displays the density of vectors in the CFD mesh

$$\frac{\delta \mathcal{E}}{\delta \overline{\mathbf{u}}} = \mathcal{P}^{\dagger}(\mathcal{P}\overline{\mathbf{u}} - \overline{\mathbf{u}}_{\exp}). \tag{17}$$

The adjoint operator of $\mathcal{P}$ satisfies

$$\left\langle \mathcal{P}^{\dagger}\Delta\mathbf{u}, \delta\overline{\mathbf{u}} \right\rangle = \langle \Delta\mathbf{u}, \mathcal{P}\delta\overline{\mathbf{u}} \rangle_M \text{ for all } \Delta\mathbf{u} \text{ and } \delta\overline{\mathbf{u}}. \tag{18}$$

For the $\mathcal{P}$ chosen in this study, the adjoint operator acting on $\Delta\mathbf{u}$ can be shown to be

$$\mathcal{P}^{\dagger}\Delta\mathbf{u} = \sum_{j=1}^{M} \Delta\mathbf{u}_j \phi_j(x), \tag{19}$$

and so the final expression for the forcing term of Eq. (8a) reduces to the following expression:

$$\frac{\delta \mathcal{E}}{\delta \overline{\mathbf{u}}} = \sum_{j=1}^{M} (\mathcal{P}_j \overline{\mathbf{u}} - \overline{\mathbf{u}}_{j,\exp})\phi_j(x). \tag{20}$$

It should be noted that this calculation need not include every PIV vector in the experimental data set. It is possible to artificially lower the resolution by omitting one out of two vectors in both the x and y directions, for example, which would increase the area of each cell by a factor of four. It is also possible to change the weight of a cell to increase its relative importance or turn it off completely to artificially decrease the PIV window. A second projection operator $\mathcal{S}$ could be introduced into Eq. (13) to account for artificially reducing the number of experimental measurements:

$$\Delta\mathbf{u} = \mathcal{S}(\overline{\mathbf{u}}_{\exp} - \mathcal{P}\overline{\mathbf{u}}). \tag{21}$$

If $\mathcal{S}$ is the identity, then all experimental data are used to compute the fitting criterion and drive the data assimilation process. This will be referred to as the full-field case. The modifications to the previous derivations for $\frac{\delta \mathcal{E}}{\delta \overline{\mathbf{u}}}$ are minor even though it will change the measurement space in Eq. (12) as the effect of $\mathcal{S}$ can be incorporated into $\phi_j(x)$ of the original smoothing function $\mathcal{P}$.

The smoothing procedure can be interpreted as a way to compute the pseudo-inverse of the projection operator $\mathcal{P}$. For this setup, the numerical data must be projected down onto the experimental data subspace by $\mathcal{P}$ to compute the discrepancy field but the adjoint solution needs to be forced by a quantity defined on the same subspace as the original numerical data. The method outlined in the previous paragraph is one way to perform this pseudo-inverse operation. Alternatives, which were not implemented in this study, include enforcing global smoothness constraints or including regularization terms in Eq. (7). Failure to include the smoothing procedure results in forcing the adjoint equations by Dirac delta functions located at the center of each cell and the

resulting forcing fields are no longer smooth. The algorithm, furthermore, will only attempt to match the velocities at discrete points where the PIV vectors are located, and this can lead to spurious recirculation bubbles or other unphysical flow structures appearing in the assimilated flow field between measurement points.

An additional complication with the experimental setup is being able to obtain PIV measurements on both sides of the airfoil. The laser sheet can only illuminate one side of the body, which casts a shadow on the opposite side. Since the airfoil is symmetric and oriented at a zero angle of attack, this problem can be overcome by computing the discrepancy velocity field for cells above the centerline only and then reflecting the measurements to the cells below the centerline. A non-symmetric airfoil or an airfoil at a non-zero angle of attack would require experimental flow measurements on both sides of the body (see Suzuki et al. 2009a, b). The streamwise discrepancy velocity is unaltered while the direction of the transverse discrepancy velocity is reversed after performing this reflection. Both the smoothing procedure, which needs to be applied to all cases in this article, and the discrepancy field reflection account for two of the modifications referred to in Sect. 3.

4.3 Adjoint looping

The block diagram in Fig. 4 illustrates the adjoint looping procedure used to iteratively update the forcing $\mathbf{f}_s$. The process is terminated once the new steepest descent direction no longer results in a reduction of the fitting criterion. The data assimilation begins with an initial guess $\mathbf{f}_0 = 0$, which is used to solve the RANS equations given by Eq. (4). Next, the discrepancy field is computed using
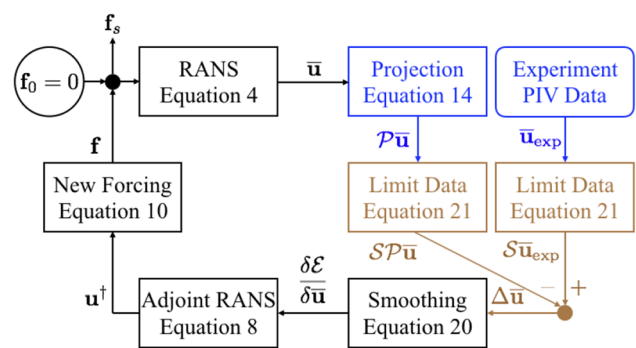
Fig. 4 A block diagram representation of the adjoint looping procedure used to determine the optimal forcing $\mathbf{f}_s$ which will minimize the fitting criterion. The steps and quantities in *black* are in the subspace of the numerical data, while *blue* indicates the subspace of the experiment and *brown* represents the subspace of fewer measurements after applying the projection operator $\mathcal{S}$

Eq. (21) once the relevant projection operators have been applied to $\bar{\mathbf{u}}$ and $\bar{\mathbf{u}}_{\text{exp}}$, and these values are smoothed over all mesh points in a grid cell as explained earlier. Finally, the adjoint equations are solved to find the steepest descent direction that updates the forcing.

A problem that arises in this procedure is finding a base flow solution, or a solution to the steady Navier–Stokes equations with zero forcing, at the experimental Reynolds number. This is due to the fact that the critical Reynolds number, or the Reynolds number at which the base flow undergoes a supercritical Hopf bifurcation and becomes unstable, tends to be rather low (for this geometry it is $\mathcal{O}(10^3)$). The base flows at higher Reynolds numbers, as a result, tend to be unphysical and require exceedingly high numerical precision to compute (Sipp et al. 2010).

The third modification that is necessary to account for a relatively high Reynolds number is to conduct the data assimilation process starting at a lower Reynolds number. A Reynolds number of $Re = 2500$ is chosen since it is possible to compute a base flow solution for the idealized airfoil. Once the fitting criterion is minimized for $Re = 2500$, the Reynolds number is increased in increments of approximately 2500 to reach the experimental Reynolds number of Re=13,500. The choice of these parameters does not have an impact on the results as long as the Reynolds number is not increased too quickly.

5 Results and discussion

The data assimilation procedure is conducted for the idealized airfoil using the time-averaged PIV data set. The resolution and domain size are artificially altered through

the use of projection operators to determine the minimum number of measurements for successful mean flow reconstruction.

5.1 Base flow and PIV results

Before introducing the results of the data assimilation algorithm, it is useful to compare the time-averaged PIV flow with the base flow solution at $Re = 2500$, which is used as an initial guess to the optimization. Figure 5 includes contour plots of the streamwise velocity for both the mean flow and the base flow as well as contour plots of vorticity overlaid with streamlines of the flow. Note that experimental data below the airfoil are not available since the laser sheet does not illuminate both sides of the airfoil.

The most significant difference between the two fields is the streamwise length of the recirculation bubble, which for the base flow extends multiple chord lengths downstream of the airfoil. This is consistent with the observation that recirculation bubble lengths for base flows around closed bodies tend to scale linearly with the Reynolds number even when the flows become unstable as reported by Zielinska et al. (1997) for the cylinder wake.

5.2 Full-field information

The data assimilation procedure is first conducted using full-field knowledge or setting $\mathcal{S}$ equal to the identity matrix in Eq. (21). This is referred to as Case A. Figure 6 compares the mean flow with the data-assimilated flow in a fashion similar to the comparison between the mean and base flows.

Overall, very good agreement is observed between both mean velocity fields, particularly when comparing the contours of the streamwise velocity. The size, streamwise position, and shape of the recirculation bubble are
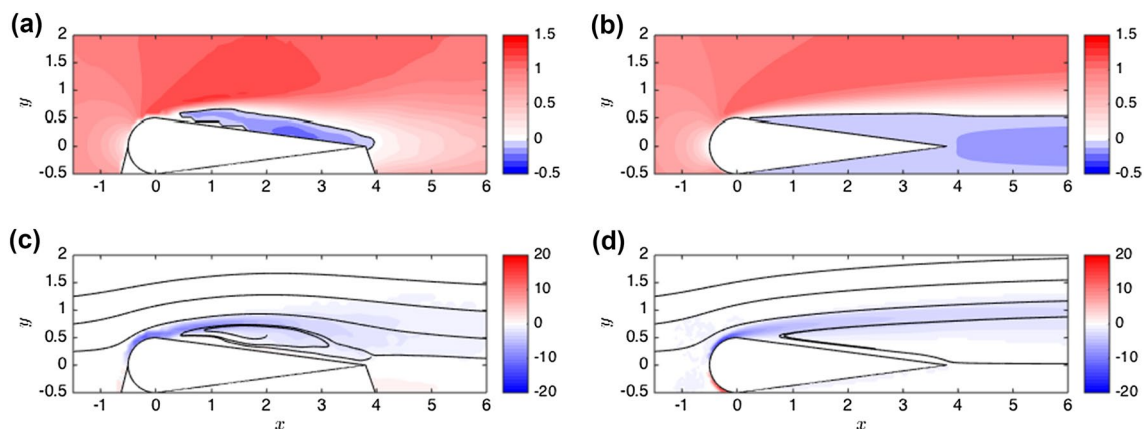
Fig. 5 Base flow: a comparison between the mean flow (*left*) obtained from PIV and the base flow for $Re = 2500$ (*right*) around the idealized airfoil. The flows are visualized by contours of the streamwise velocity (*top*) and vorticity (*bottom*). The *black contour* in (**a**) and (**b**) corresponds to $\bar{u} = 0$ and streamlines are included in (**c**) and (**d**)
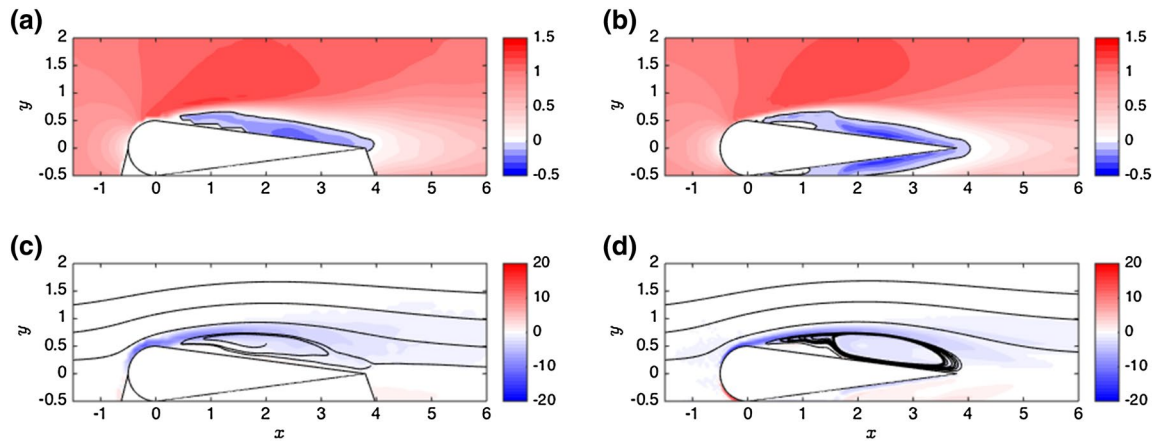
Fig. 6 Case A results: a comparison between the mean flow (*left*) obtained from PIV and the data-assimilated flow (*right*) around the idealized airfoil. The flows are visualized by contours of the stream- wise velocity (*top*) and vorticity (*bottom*). The *black contour* in (**a**) and (**b**) corresponds to $\bar{u} = 0$ and streamlines are included in (**c**) and (**d**).

also captured quite well although there is a slight discrepancy with the center of the bubble. It is difficult to determine this location from the PIV data due to the lack of spatial resolution in the transverse direction. The $\bar{u} = 0$ contour does not, however, match near the front of the bubble and this discrepancy is discussed in Sect. 6.

To quantify how well the data assimilation recovers the mean flow, the L_2-difference is computed between the PIV velocity field $\bar{\mathbf{u}}_{\text{exp}}$ and the assimilated flow $\bar{\mathbf{u}}$ interpolated onto the PIV grid, which will be referred to as $\bar{\mathbf{u}}_{\text{int}}$. This difference is then normalized by the L_2-norm of the PIV velocity field and is referred to as the experimental mismatch. It can be written as

$$\mathcal{R} = \frac{\left\| \bar{\mathbf{u}}_{\text{exp}} - \bar{\mathbf{u}}_{\text{int}} \right\|}{\left\| \bar{\mathbf{u}}_{\text{exp}} \right\|}. \tag{22}$$

$\mathcal{R}$ is computed on the domain $x \in [-1.5, 6]$; $y \in [0, 2]$. This definition is constant across all cases studied even if the experimental data are artificially limited.

Another quantity of interest is the decrease of the cost function for each case investigated. A bad cost function is one which leads to a large decrease in the fitting criterion without reducing the experimental mismatch. To determine the success of a given cost function, its final value $\mathcal{E}_{\text{end}}$ is normalized by the initial value computed using the base flow $\mathcal{E}_0$. This ratio can be compared to the corresponding decrease of the experimental mismatch

$$\mathcal{R}_r = \frac{\mathcal{R}_{\text{end}}}{\mathcal{R}_0}, \quad \mathcal{E}_r = \sqrt{\frac{\mathcal{E}_{\text{end}}}{\mathcal{E}_0}}, \tag{23}$$

where $\mathcal{R}_0$ is computed using the base flow. The square root of the cost function ratio is necessary since $\mathcal{E}_r$ is based on $\| \cdot \|^2$ while $\mathcal{R}_r$ is based on $\| \cdot \|$. If $\mathcal{E}_r << \mathcal{R}_r$, the cost

Table 1 Summary of the data assimilation results

Case	Description	Number of meas. points	$\mathcal{R}_{\text{end}}$	$\mathcal{R}_r$	$\mathcal{E}_r$
A	Full field	3669	6.76%	25.9%	40.1%
B	Half resolution	934	8.17%	31.3%	36.1%
C	Third resolution	416	11.5%	44.1%	47.2%
D	Small field	1765	7.73%	26.1%	48.6%
E	Half small	465	8.79%	33.7%	39.9%

function is not effectively reducing the mean velocity field discrepancy.

For the full-field case (Case A), the cost functional decreases by $\mathcal{E}_r = 40.1\%$ and the experimental mismatch is $\mathcal{R}_{\text{end}} = 6.76\%$, down from an initial value of 26.1% when the experimental mismatch is calculated with respect to the base flow ($\mathcal{R}_r = 25.9\%$). These results are summarized in Table 1. It is worth noting that the initial mismatch is misleadingly low since the PIV data does not extend far downstream ($x \leq 6$) and the base flow is at a Reynolds number lower than that of the experiment. The real base flow, which is very difficult to compute, would be even more unphysical than the current one. The region of the flow behind the airfoil is where the largest discrepancies are concentrated since the recirculation region is overestimated by the base flow.

The last result for the full-field information case is to compare the forcing of the assimilated flow with that of the experiment. Since the algorithm can only capture the solenoidal part of the forcing, the curl of the forcing field $\nabla \times \mathbf{f}$ is presented in Fig. 7 instead of $\mathbf{f}_x$ and $\mathbf{f}_y$ individually. Computing this quantity from the experiment, however, requires two spatial derivatives of the Reynolds
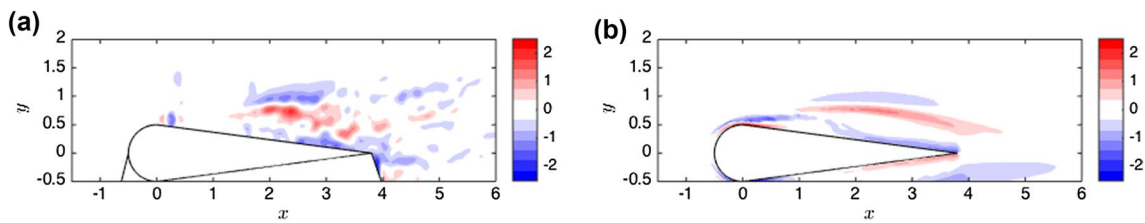
Fig. 7 Case A results: a comparison between curl of the forcing $\nabla \times \mathbf{f}$ for the experimental velocity field (*left*) and assimilated velocity field (*right*)

stresses. A circular averaging or 'disk' filter in Matlab was used to smooth out the Reynolds stresses before computing gradients. No smoothing was necessary for the assimilated forcing.

There is good agreement between the structure of the forcing fields (note that it is actually the curl which has been reproduced in Fig. 7 but it shall be referred to as forcing for convenience) and the magnitudes of the peaks in the shear layer. It should be noted that complete agreement is not expected since the PIV data are not ideal and there is underlying noise in the measurements. The assimilation is also able to reconstruct the forcing at the leading edge of the airfoil, information which is not available from the experimental data since there are too few PIV measurements to resolve the high Reynolds stress gradients in this region. It is striking to observe how clean the assimilated forcing looks in comparison to the experimental one. The results reinforce how one of the primary motivations of data assimilation is to remove noise and produce more highly resolved flow quantities. Assimilated fields are particularly useful at reproducing fields which need to be differentiated since experimental derivatives tend to amplify the underlying noise in a measurement.

5.3 Minimum PIV resolution

The full-field assimilation procedure uses 3669 PIV vectors. Using the projection operator $\mathcal{S}$, it is possible to investigate the effect of decreasing the full-field resolution and its impact on the assimilated flow fields. Three different resolutions are tested by removing points in both the streamwise and transverse directions in a fixed pattern. To obtain an experimental field with half the resolution of the original one, for example, every other point is removed in both the streamwise and transverse directions. This process is repeated to obtain data sets with one third of the original resolution as well as one fourth of the original resolution.

Results for the assimilated flows at the lower resolutions are presented below in Fig. 8 for the one-half and one-third case (Case B and Case C, respectively). The one-fourth resolution case failed to converge at higher Reynolds numbers, so the results have been omitted.

The experimental mismatch calculations are summarized in Table 1. It is evident that there is a noticeable decrease in the assimilated flow field quality for the one-third case in comparison to the one-half case. The recirculation bubble is not smooth and the velocity contours have significantly deviated from the mean. This is largely due to the fact that the cell size over which the smoothing procedure is applied is nine times larger than it was in the full-field case. As a result, the range of velocities in a given cell can be quite large, particularly when the cell encompasses measurements above the shear layer where the velocity is close to the free stream and below the shear layer where it is close to zero. Nevertheless, the one-third case correctly reproduces the main features of the mean field, particularly the length and height of the recirculation bubble. It is also manages to recover an acceptable approximation of the forcing field as shown in Fig. 9.

It is important to note that the discrepancy field calculated by Eq. (21) is not smooth since all the mesh points in a given cell have the same value. This field becomes decreasingly smooth as the resolution of the PIV data decreases since the cell size over which the averaging operator applies increases. Despite how discontinuous this smoothing may be, the solution to the adjoint equations, and consequently the recovered forcing field, is quite smooth as seen in Fig. 9. It is worth noting, however, that the one-third case does not accurately locate the maximum contours of the forcing. The maximum positive contour in the forcing field, for example, is located further downstream when compared to the full-field and one-half cases.

5.4 Minimum field of view

Another way to reduce the number of experimental measurements is to truncate the PIV field of view. In the interest of brevity, only the smallest field of view for which there was no major sacrifice in data assimilation quality is presented. The domain was limited such that PIV points outside the range $-1.5 < x < 6.0$, $0 < y < 1.5$ were excluded. The results are quantified in Table 1 as Case D.

The trends observed in Table 1 are consistent with intuition—the more experimental measurements available,
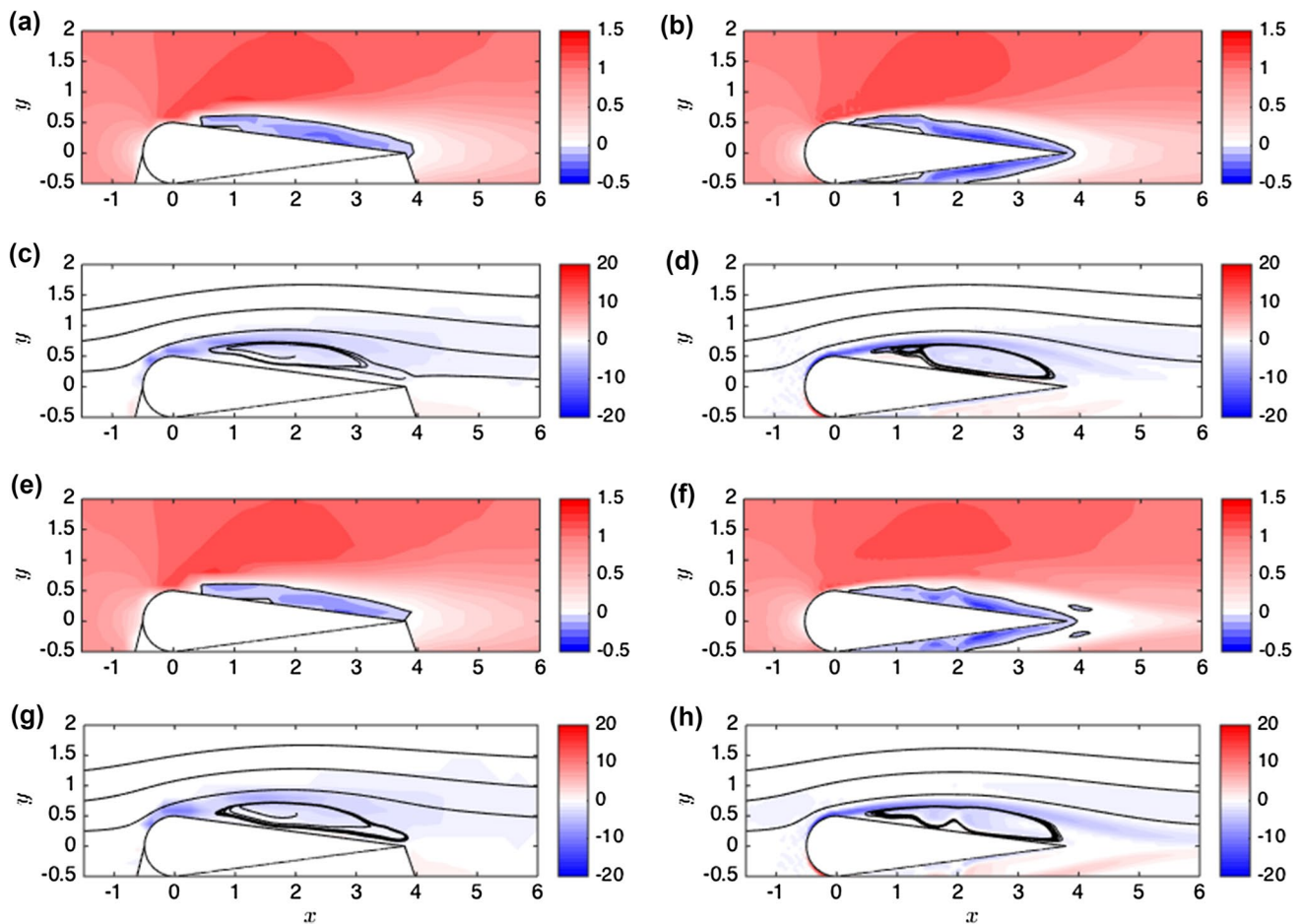
Fig. 8 Case B and C results: the projected (experimental) mean flows are shown by (**a**) and (**c**) for the one-half resolution case (Case B) and (**e**) and (**g**) for the one-third resolution case (Case C). The corresponding assimilated flows are (**b**) and (**d**) for the one-half resolution case and (**f**) and (**h**) for the one-third resolution case. The flows are visualized by *contours* of the streamwise velocity and vorticity

the closer the assimilated flow field is to the experiment. A promising result, however, is that decreasing the resolution or field of view of the reference data set does not significantly affect the assimilated flow field quality up to a point. As mentioned earlier, there is a significant difference between the one-half and one-third resolution cases. It is difficult to know a priori how much resolution is needed as this depends largely on the flow structures in the mean flow. The one-third resolution case struggles to resolve the shear layer while the one-fourth resolution case (not shown due to convergence problems) used grid cells which were too large to resolve the mean recirculation bubble. Predictive tools, which can identify the location and size of these features, would be helpful to determine the experimental measurements which are necessary; they will be the subject of future work.

To reduce the number of measurements further, it is possible to combine the two approaches mentioned so far by truncating the field of view and reducing the resolution of

the reference data. Combining the small domain with the half resolution case decreases the number of points by approximately a factor of ten and is referred to as Case E. A comparison of the assimilated flow using the approach with the mean flow is presented in Fig. 10. The forcing fields are also displayed in Fig. 11 to demonstrate that reduction in points by a factor of approximately ten has only a minimal impact on the results.

Data assimilation works well within the current framework and reproduces smooth velocity and forcing fields. There are, however, limitations to this analysis that are worth mentioning. To begin with, the algorithm needs several hundred spatial measurements for successful mean flow reconstruction. It is certainly possible to devise projection operators which divide the experimental data into variable cell sizes so that regions of the flow with high velocity gradients are more highly resolved than other regions. Experimental data, especially from PIV, is typically not collected on non-uniform grids so
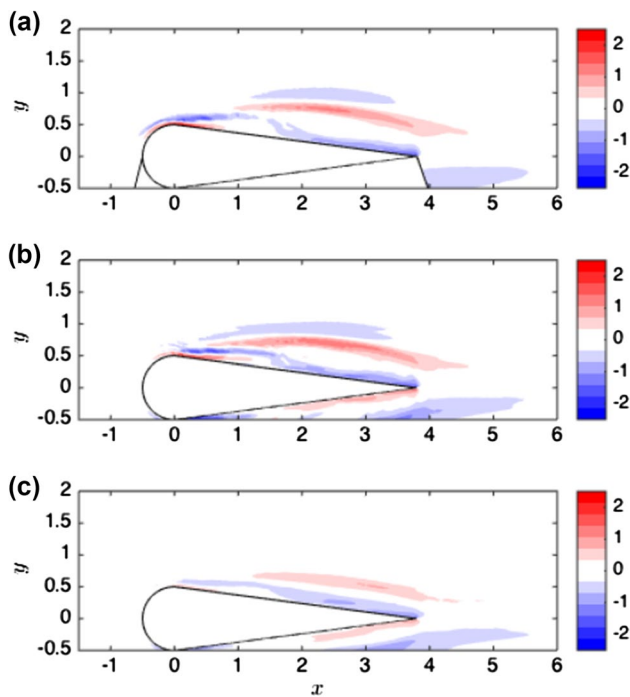
Fig. 9 Case B and C results: a comparison between the curl of the forcing $\nabla \times \mathbf{f}$ for all three resolutions. The full-field assimilated forcing (**a**) is reproduced from Fig. 7 to compare against the forcing from the one-half resolution case (**b**) and the one-third resolution case (**c**)

this study limits the approach to projection operators which could be reproduced by simply altering the experimental parameters. There are also temporal averaging requirements since a sufficient number of measurements is needed to obtain converged statistics. For bluff bodies, this might require some a priori knowledge of the shedding frequency, while for wall-bounded turbulent flows it is necessary to determine u_τ and the eddy turnover time.
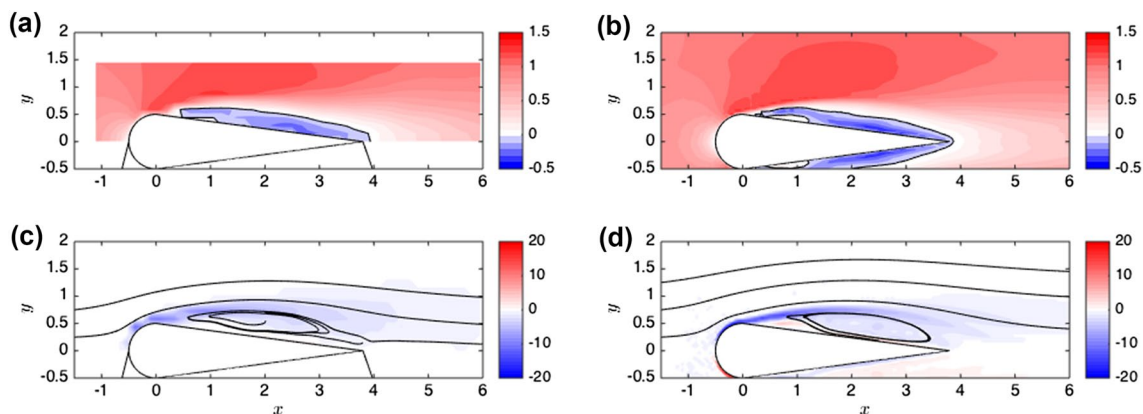
6 Residual discrepancy

An assessment of the remaining discrepancy between the experimental and assimilated mean velocity fields is discussed in this section. The influence of three-dimensionality of the flow is considered as well as elements of the data assimilation framework.

6.1 Three dimensionality

To quantify three-dimensional effects, the divergence of the experimental velocity is calculated. If the 2D continuity equation is not satisfied then this constraint may not be appropriate. The same calculation is performed for the data-assimilated flow field from Case A which is constrained to be divergence-free in 2D. This field is interpolated onto the PIV grid to determine what the permissible range of values would be for a flow field obeying the 2D continuity equation. Figure 12 compares $\nabla \cdot \overline{\mathbf{u}}$ for the experiment and the assimilation.

The plots indicate that $\nabla \cdot \overline{\mathbf{u}}$ is nearly zero everywhere except near the leading edge. Since $\nabla \cdot \overline{\mathbf{u}}$ is also high for the data-assimilated field at the leading edge it is reasonable to suggest that the grid resolution is not sufficient to capture the large velocity gradients in this region of the flow.

As mentioned earlier, 3D effects would be compensated for by the unknown momentum forcing term $\mathbf{f}_s$. It is clear from Sect. 5 that $\mathbf{f}_s$ does not deviate significantly from the experiment. This means it is not being corrupted by spanwise velocity gradients which have been assumed to be zero. From these two observations, the role of 3D effects is negligible and does not play a significant role in the residual discrepancy.

Fig. 10 Case E results: a comparison between the projected (experimental) mean flow (*left*) obtained from PIV and the data-assimilated flow (*right*) around the idealized airfoil. The flows are visualized by contours of the streamwise velocity (*top*) and vorticity (*bottom*). The black contour in (**a**) and (**b**) corresponds to $\overline{u} = 0$ and streamlines are included in (**c**) and (**d**)

6.2 Model simplifications

There are two modeling simplifications in the data assimilation framework: boundary conditions and the coordinate system mapping between the experiment and the simulation. The boundary conditions for the simulation are very general and do not take into account the blockage ratio, which for this experimental configuration is approximately 7%. The PIV domain, furthermore, does not extend sufficiently far in the transverse direction to observe effects from the wall. Since there is excellent agreement between the assimilation and experiment outside the recirculation bubble of the airfoil, the effects of the boundary conditions and blockage ratio are negligible.

The most significant source of the residual discrepancy, as seen in Fig. 13, is the mapping between the experiment and the simulation. The largest contribution comes from several points along the leading edge of the airfoil while there are smaller contributions in parts of the recirculation

bubble as well as the airfoil boundary near the trailing edge. There are two reasons to account for these results. First, it is difficult to precisely determine the airfoil location from the PIV data and second, there are imperfections in the airfoil shape which is modeled as a cylinder followed by a wedge of half angle 7.5°. Consequently, great care is taken to approximate the location and shape of the airfoil to minimize the uncertainty that arises from the mapping.

6.3 Choice of cost function

This study considers simple cost functions where the objective is to minimize the discrepancy at all PIV points subject to the incompressible RANS equations. No regularization parameters are introduced and all the PIV vectors are weighted equally. While effective, the major drawback to such an approach is that low-velocity regions of the flow are treated as less important. The discrepancy between the velocity in the recirculation bubble, for example, is low
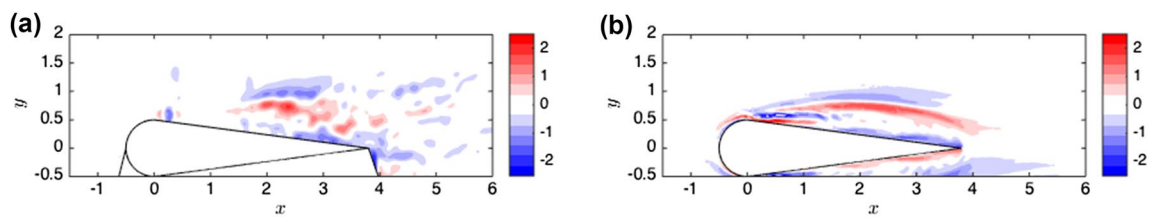
Fig. 11 Case E results: a comparison between the curl of the forcing $\nabla \times \mathbf{f}$ for the experimental velocity field (*left*) and the assimilated velocity field (*right*). Note that the projection operator has not been applied to the experimental forcing since this would exacerbate the level of noise
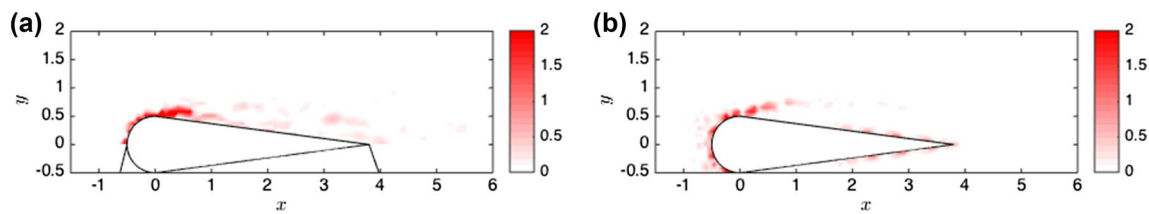
Fig. 12 *Contours* of $\nabla \cdot \overline{\mathbf{u}}$ computed for the experiment (*left*) and the data-assimilated field (*right*) which is constrained to satisfy the 2D continuity equation
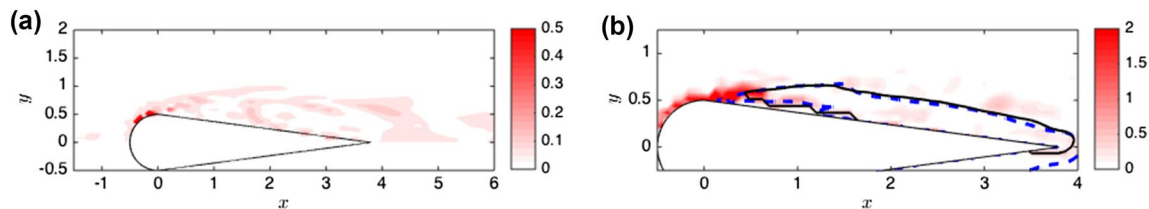
Fig. 13 Residual discrepancy between the experimental and assimilated mean velocity fields for Case A. The *left-hand side* quantifies the magnitude of the velocity discrepancy while the *right-hand side*

shows the $\overline{u} = 0$ *contour* for the experiment (*black*, *solid line*) and assimilation (*blue*, *dotted line*) overlaid with *contours* of $\nabla \cdot \overline{\mathbf{u}}$

even for the base flow case. Consequently, the procedure is biased towards high-speed regions of the flow and it tries to correct these regions first.

The right-hand plot of Fig. 13 compares the $\overline{u} = 0$ contour of Case A with its counterpart from the PIV data. The plot also includes contours of $\nabla \cdot \overline{\mathbf{u}}$ to emphasize that three dimensionality is not the leading cause of the discrepancy. There are small deviations along the entire contour which could be addressed by modifying the cost function to penalize these discrepancies to a higher degree.

7 Measurement uncertainty

It is useful to give estimates of various sources of uncertainty in the experiment to quantify the precision of data entering the data assimilation algorithm. This section identifies these sources and offers a suggestion for incorporating them into the framework. Measurements with higher uncertainty are deemed less trustworthy and the data assimilation algorithm would take this into account by converging to the experimental measurement within the uncertainty bound instead of the exact experimental value.

The uncertainty can be decomposed into several components, each associated with a particular aspect of the experimental setup, data acquisition, or data processing. These include experimental setup assumptions ($\mathbf{E}_{\text{setup}}$), the sequential-correlation computations of the PIV velocity vectors ($\mathbf{E}_{\text{PIV}}$), the mapping from the experimental to the numerical coordinate system ($\mathbf{E}_{\text{map}}$), and the variance error from the averaging procedure ($\mathbf{E}_{\text{var}}$). The combined measurement error $\mathbf{E}$ is taken to be the sum of these components, i.e.,

$$\mathbf{E} = \mathbf{E}_{\text{setup}} + \mathbf{E}_{\text{PIV}} + \mathbf{E}_{\text{map}} + \mathbf{E}_{\text{var}}. \tag{24}$$

Two notations are adopted: $\mathbf{E}$ denotes the relative uncertainty, while $\delta\overline{\mathbf{u}}_{\text{exp}}$ stands for the absolute uncertainty of the assimilated velocity $\overline{\mathbf{u}}_{\text{exp}}$; consequently,

$$\mathbf{E} = \frac{\delta\overline{\mathbf{u}}_{\text{exp}}}{\overline{\mathbf{u}}_{\text{exp}}}, \tag{25}$$

expresses the link between relative and absolute uncertainty.

7.1 Experimental setup

There are many possible sources of error due to factors related to the experimental setup that could influence a PIV measurement. Adrian and Westerweel (2011) mention factors such as calibration, timing errors, particle slip, and the optical setup. For this particular flow,

only the optical setup and calibration are considered as sources of experimental error and their contributions to the overall uncertainty are estimated.

The extraction of velocity vector fields using PIV assumes that both the laser sheet and the calibration plate are exactly coincident with the streamwise-transverse velocity plane. With a laser sheet measuring approximately 1 mm in thickness and three airfoil chords in length, the error in parallelism is estimated to be on the order of 20 arc minutes (0.33°). A non-aligned laser sheet poses a problem if a significant number of particles traverse from an illuminated area to a non-illuminated one (or vice versa) between two consecutive camera snapshots. Particles in this flow move up to 0.8 mm between two images (based on the free-stream velocity). Additionally, about 99% of the particles in the free stream are illuminated by the laser sheet during two successive snapshots. For the smallest correlation window (16×16 pixels), there are thus between 10 and 25 particles, resulting in a 10% to 25% probability for one particle to leave (or enter into) the illuminated region and to influence the sharpness of the correlation peak. These estimates show that errors stemming from misalignment of the laser sheet are expected to be very small based on the results of Wieneke (2015), who quantified the error in pixels due to out-of-plane motion, and negligible in view of other sources of error, as detailed below.

The calibration plate, on the other hand, may also be tilted by up to 20 arc minutes with respect to the streamwise–spanwise plane. This angle uncertainty translates directly into a 0.01% measurement uncertainty and so a total error of about 0.01% in the measured velocity should be attributed to misalignment factors in the experimental setup.

The cameras are calibrated using a calibration plate placed at the laser sheet location. This plate consists of a grid of known dimensions which translates between the camera-sensors' coordinate system and the physical domain. Typically, this mapping has a finite number of degrees of freedom (e.g., a pinhole model with eight parameters) and is computed using a least-squares fit. The non-zero RMS error, i.e., the residual of this fit, is the uncertainty associated with camera calibration. In this study, this uncertainty is between 0.45 and 0.5 pixels for both cameras. Thus, when the velocity is evaluated based on the position of the correlation peak, it must be kept in mind that the position of this peak is known, on average, up to half a pixel. Based on the sources outlined in this section and the size of the correlation window, the instantaneous velocity is estimated to be within an uncertainty bound of 2% so $\delta\mathbf{u}_{\text{set-up}} = 0.02\mathbf{u}$.

7.2 PIV correlation peak position

The bias error stemming from the PIV correlation computations is a source of uncertainty that can be readily estimated. Even though the displacement of the correlation peak can be induced by a variety of physical phenomena, only effects due to particle size and velocity variations within the interrogation window will be considered for simplicity. The bias ϵ of the correlation peak (in pixels) can be expressed explicitly for the case of simple shear and under the assumption of a Gaussian correlation function. With an interrogation window of dimension D_I, image magnification M, time interval Δt between images, and particles of identical diameter d_τ (see Westerweel 2008), then

$$\frac{\epsilon}{D_I} = \frac{1}{F_I(\mathbf{s}_D)} \left(\frac{d_\tau^2}{8D_I^2} + \frac{(M|\Delta u|\Delta t)^2}{12D_I^2} \right). \tag{26}$$

The local variation of the velocity field $|\Delta u|$ is approximated as

$$|\Delta u| \sim \left| \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{x}} \right| \cdot L, \tag{27}$$

where L is a typical dimension of the interrogation volume. $F_I(s_D)$ is the in-plane loss of correlation function, which depends on $\mathbf{s}_D$, the position of the correlation peak within the interrogation window. $F_I(s_D)$ can be written as

$$F_I(\mathbf{s}_D) = \begin{cases} \left(1 - \dfrac{|s_x|}{D_I} \right)\left(1 - \dfrac{|s_y|}{D_I} \right) & \text{for } |s_x| < D_I, \, |s_y| < D_I, \\ 0 & \text{otherwise.} \end{cases} \tag{28}$$

The velocity bias error $\delta\mathbf{u}_{\text{PIV}}$ due to PIV correlation peaks may be expressed in terms of the correlation peak displacement, leading to the quantification of the PIV-measurement uncertainty as

$$\delta\mathbf{u}_{\text{PIV}} = \frac{\epsilon}{M \, \Delta t}. \tag{29}$$

Due to Eq. (27), this uncertainty varies spatially, since it depends on the velocity gradients. Wilson and Smith (2013) point out that the presence of shear, or large velocity gradients, broadens the correlation peak and leads to non-uniform particle image densities. The resulting effect on the mean uncertainty for this experiment, therefore, can be greater than 0.1 pixels. The contribution for this source of uncertainty on the instantaneous velocity is included in the calculation performed in Eq. (31).

7.3 Coordinate system bias

As mentioned in Sect. 6, the mapping from the experimental to the numerical coordinate system is not perfect. The center of the airfoil is estimated in the calibrated plane using the background snapshot in which the airfoil is illuminated. This error will change the velocity at a given point, depending on the velocity gradient. Measurement points near the leading edge where gradients are highest would be considered less trustworthy. Assuming the associated error $\delta\mathbf{u}_{\text{map}}$ is of the order of the velocity gradient multiplied by the position error C_e, the resulting uncertainty is approximated as

$$\delta\mathbf{u}_{\text{map}} = \|\nabla\mathbf{u}\| C_e. \tag{30}$$

C_e is set to a value of 0.02 since there are approximately 50 vectors along the chord of the airfoil and the position of the airfoil is accurate to within one vector.

7.4 Variance error during averaging

The mean velocity is computed by taking the average over all measurement snapshots. Because of a limited number of snapshots, an error arises which can be estimated by considering the convergence of the mean quantities. Sciacchitano and Wieneke (2016) derive the uncertainty of the mean velocity as

$$\mathbf{E}_{\text{var}} = \frac{\sigma_u}{\sqrt{N}}, \tag{31}$$

where σ_u contains the true velocity fluctuations and the measurement errors and N is the number of samples. This expression does not take into account the systematic uncertainties mentioned earlier. These are included when combining all measurement uncertainties using the method outlined by Wilson and Smith (2013).

7.5 Combined measurement uncertainties

The combined measurement uncertainties are propagated into the mean using Eq. (31) from Sciacchitano and Wieneke (2016). The combined uncertainty for the mean streamwise velocity $\bar{u}$ is greater than $\bar{v}$ as seen in Fig. 14. The largest uncertainties, which are below 0.04, are concentrated near the cylindrical leading edge where the velocity gradients are highest. It should be noted that the relative uncertainty, which has not been plotted since points where the mean velocity is close to zero overwhelm the contours, is highest in the reverse flow region. While the absolute uncertainty is close to zero in this region, the relative uncertainty can exceed 50%.
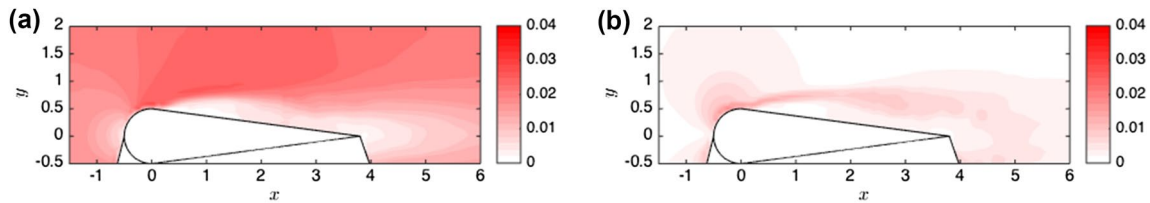
Fig. 14 Absolute uncertainty of the experimental PIV measurements normalized by the unit inlet velocity. The uncertainty of the mean streamwise velocity $\sigma_{\overline{u}}$ is *plotted* on the *left* while the transverse velocity counterpart $\sigma_{\overline{v}}$ is *plotted* on the *right*

Although it was not implemented for this particular study, the uncertainty bounds on the data could, in the future, be incorporated into the fitting criterion. Defining a general, real-valued function N that accounts for measurement uncertainty, the fitting criterion can be recast as

$$\mathcal{E}(\overline{\mathbf{u}}(\mathbf{f}_s)) = \mathsf{N}(\Delta\mathbf{u}). \tag{32}$$

In addition, it is assumed that the total measurement error (from all known sources) can be bounded by $\mathbf{E}$. The derivative of N with respect to measurement discrepancies can then be defined as

$$\mathsf{N}'(\mathbf{x}) = \begin{cases} \mathbf{x} + \mathbf{E} & \text{for} \quad \mathbf{x} < -\mathbf{E} \\ \mathbf{0} & \text{for} \quad |\mathbf{x}| \le \mathbf{E} \\ \mathbf{x} - \mathbf{E} & \text{for} \quad \mathbf{E} < \mathbf{x} \end{cases} \quad \text{and} \quad \mathsf{N}(0) = 0. \tag{33}$$

With this definition, the function N is constant for all measurement discrepancies $\Delta\mathbf{u}$ if $|\Delta\mathbf{u}| \le \mathbf{E}$. The modified fitting criterion in Eq. (32) will direct the algorithm to produce a solution that matches the measurements up to their uncertainties. Eq. (33) is a hard-bound approach to account for uncertainty in the data assimilation. Alternatives include the implementation of a soft-bound expression for the fitting criterion or a standard Tikhonov regularization approach combined with weight functions proportional to uncertainty bounds (see Bukshtynov et al. 2011; Flemming 2011). The modified fitting criterion impacts the optimization algorithm, and consequently the following adjustment needs to be made

$$\frac{\delta\mathcal{E}}{\delta\overline{\mathbf{u}}} = \mathsf{N}'(\Delta\mathbf{u}). \tag{34}$$

Since N only appears as its derivative, it seems more suitable to define N by its derivative, as in Eq. (33). The above expression would replace the right-hand side of Eq. (8a) during the optimization procedure to account for measurement uncertainty.

8 Conclusions

The data assimilation framework based on Foures et al. (2014) is extended to recover the mean flow and unknown momentum forcing around an idealized airfoil at a Reynolds number of Re=13,500. The experimental data set originates from time-resolved PIV data on a uniform Cartesian grid with a spatial resolution far lower than that of the numerical simulation. It is possible, nevertheless, to compute an assimilated flow field which is in very good agreement with the experiment. The recovered forcing field is also in good agreement with the experimental forcing and contains dramatically less noise since the experimental forcing is computed by differentiating Reynolds stress fields.

Three modifications to the original framework are necessary for practical implementation of the algorithm. First, it is necessary to use a smoothing procedure to ensure that the adjoint equations are being forced at all mesh points instead of just those which lie closest to the experimental grid points. Second, discrepancy velocity measurements are computed above the centerline only and are reflected to account for the shadow cast by the airfoil. Finally, it is necessary to begin with a base flow computed at a lower Reynolds number than that of the experiment due to the difficulty of computing a base flow at Reynolds numbers much higher than the critical one. Once the data assimilation process is complete for a lower Reynolds number, it can be incrementally increased until it matches the experimental conditions.

As expected, the full-field case where all PIV vectors are used to guide the mean flow reconstruction yields the best reconstructed fields. It is encouraging, nevertheless, that truncating the PIV field of view or decreasing the resolution does not significantly impact the quality of the results. Improved smoothing procedures or weighting measurement points where important mean flow features such as the recirculation bubble are present in the flow could help reduce the number of necessary points further. As alluded to in Sect. 5, future work is necessary for a priori guidance on where it is important to obtain experimental measurements and how well in space they should be resolved. The quality of the results could also be improved by implementing bounds on how close the numerical values match the experiment via the fitting criterion to account for measurement uncertainty.

One of two future considerations is to account for the out-of-plane velocities which are not captured in the measurements or the processed data. While the 2D

incompressible RANS equations seem justified for this flow, an algorithm in the future could account for out-of-plane velocity gradients by introducing an additional compliance parameter $d = -\partial_z w = \partial_x u + \partial_y v$ and replacing the divergence-related term in Eq. (6) by $\langle p^\dagger, \nabla \cdot \bar{\mathbf{u}} - d \rangle$. For 2D flow ($d = 0$) this expression reverts back to the previous augmented Lagrangian given by Eq. (6). The second would be to compare the predictions of a RANS simulation with a turbulence model to the experiment and assimilated results. The degree of success is dependent on the model chosen so there is not a unique comparison between traditional RANS and the approach outlined in this study.

Progress on some of the suggested extensions is currently underway to improve the algorithm for mean flow reconstruction. Even with the relatively simple approach outlined in this paper both the reconstructed mean velocity field and forcing fields agree quite closely and are dramatically smoother than their experimental counterparts. The original experimental data set, furthermore, contains more information than necessary to reproduce the main flow features. Future work would be to extend the framework to three-dimensional flow fields which are typically more limited by a laser's ability to illuminate a control volume. Data assimilation could thus be a worthwhile tool for the experimentalist to recover more information about the measured flow.

References

Adrian RJ (1991) Particle-imaging techniques for experimental fluid mechanics. Annu Rev Fluid Mech 23:261–304

Adrian RJ, Westerweel J (2011) Particle image velocimetry. Cambridge University, New York

Bewley TR, Moin P, Temam R (2001) DNS-based predictive control of turbulence: an optimal benchmark for feedback algorithms. J Fluid Mech 447:179–225

Bukshtynov V, Volkov O, Protas B (2011) On optimal reconstruction of constitutive relations. Physica D 240(16):1228–1244

Flemming J (2011) Generalized Tikhonov regularization: Basic theory and comprehensive results on convergence rates. PhD thesis, Techn Univ Chemnitz

Foures DPG, Dovetta N, Sipp D, Schmid PJ (2014) A data-assimilation method for Reynolds-averaged Navier-Stokes-driven mean flow reconstruction. J Fluid Mech 759:404–431

Gharib M (1983) The effect of flow oscillations on cavity drag, and a technique for their control. PhD thesis, California Institute of Technology

Giannetti F, Luchini P (2007) Structural sensitivity analysis of the first instability of the cylinder wake. J Fluid Mech 581:167–197

Gonzalez C, Bruhat JF, Wainfan B, McKeon BJ (2010) Control of laminar separation on an idealized airfoil using periodic dynamic roughness actuation. In 63rd Meeting of the American Physical Society Division of Fluid Dynamics. Gallery of Fluid Motion, Poster 26

Gronskis A, Heitz D, Mémin E (2013) Inflow and initial conditions for direct numerical simulation based on adjoint data assimilation. J Comput Phys 242:480–497

Kim J, Bewley TR (2007) A linear systems approach to flow control. Annu Rev Fluid Mech 39:383–417

Kompenhans J, Raffel M, Willert C (2007) Particle image velocimetry—a pratical guide. Springer, Berlin

Lewis JM, Lakshmivarahan S, Dhall S (2006) Dynamic data assimilation: a least squares approach. In: Encyclopedia of mathematics and its application 104, vol 13. Cambridge University Press

Marquet O, Sipp D, Jacquin L (2008) Sensitivity analysis and passive control of cylinder flow. J Fluid Mech 516:221–252

Meliga P, Pujals G, Serre E (2012) Sensitivity of 2-D turbulent flow past a D-shaped cylinder using global stability. Phys Fluids 24(6):61701

Mettot C, Sipp D (2014) Quasi-laminar stability and sensitivity analyses for turbulent flows: Prediction of low-frequency unsteadiness and passive control. Phys Fluids 26(4):45112

Mons V, Chassaing JC, Gomez T, Sagaut P (2016) Reconstruction of unsteady viscous flows using data assimilation schemes. J Comput Phys 316:255–280

Nisugi K, Hayase T, Shirai A (2004) Fundamental study of hybrid wind tunnel integrating numerical simulation and experiment in analysis of flow field. JSME Int J 47:593–604

Polak E, Ribière G (1969) Note sur la convergence de méthodes de directions conjuguées. Revue française d'informatique et de recherche opérationnelle, série rouge 3(1):35–43

Prasad A, Williamson CHK (1997) The instability of the shear layer separating from a bluff body. J Fluid Mech 333:375–402

Sciacchitano A, Wieneke B (2016) PIV uncertainty propagation. Meas Sci Technol 27(8):1–16

Sipp D, Marquet O, Meliga P, Barbagallo A (2010) Dynamics and control of global instabilities in open-flows: a linearized approach. Appl Mech Rev 63(3):030801

Strykowski P, Sreenivasan K (1990) On the formation and suppression of vortex 'shedding' at low Reynolds numbers. J Fluid Mech 218:71–107

Suzuki T (2012) Reduced-order Kalman-filtered hybrid simulation combining particle tracking velocimetry and direct numerical simulation. J Fluid Mech 709:249–288

Suzuki T, Ji H, Yamamoto F (2009a) Unsteady PTV velocity field past an airfoil solved with DNS: Part 1. Algorithm of hybrid simulation and hybrid velocity field at $Re \approx 10^3$. Exp Fluids 47:957–976

Suzuki T, Sanse A, Mizushima T, Yamamoto F (2009b) Unsteady PTV velocity field past an airfoil solved with DNS: Part 2. Validation and application at Reynolds numbers up to $Re \leq 10^4$. Exp Fluids 47:977–994

Wallace RD, McKeon BJ (2012) Laminar separation bubble manipulation with dynamic roughness. In: Proceedings of the 6th AIAA flow control conference. AIAA 2012–2680

Westerweel J (2008) On velocity gradients in PIV interrogation. Exp Fluids 44:831–842

Westerweel J, Scarano F (2005) Universal outlier detection for PIV data. Exp Fluids 39(6):1096–1100

Wieneke B (2015) PIV uncertainty quantification from correlation statistics. Meas Sci Technol 26(7):1–10

Wilson BM, Smith BL (2013) Uncertainty on PIV mean and fluctuating velocity due to bias and random errors. Meas Sci Technol 24(3):1–15

Zielinska BJA, Goujon-Durand S, Dusek J, Wesfried JE (1997) Strongly nonlinear effect in unstable wakes. Phys Rev Lett 79:3893–3896