System identification for enhanced flow description from multi-pulse PIV

Patricia García-Caspueñas *, Stefano Discetti

Universidad Carlos III de Madrid, Aerospace Engineering Research Group, Madrid, Spain
*patricia.gcaspuenas@gmail.com

Keywords: PIV processing, data assimilation, low-order models, system identification

ABSTRACT

We propose a novel approach to obtain a time-resolved description from non-time-resolved PIV measurements. The method needs in input velocity and acceleration fields from (at least) 3-pulse PIV data, without time resolution. A sparse identification of nonlinear dynamics (based on the SINDy technique proposed by Brunton et al., 2016) is carried out to estimate the time history of coordinates in a low-dimensional space from Proper Orthogonal Decomposition (POD). The output is a reduced-order model of the dynamics which can be integrated to obtain time-resolved velocity fields using the snapshots as initial conditions. The time-resolved evolution of the pressure fields is then derived from the momentum equation. The time history of the velocity and pressure fields is obtained through a weighted Backward-Forward Integration (BFI) of the identified system. Time super-sampling can be achieved by integration between consecutive non-time-resolved realizations. In alternative, simple one-directional time integration can be used for short-time prediction starting from individual snapshots. The algorithm is validated using a synthetic test case from 2D Direct Numerical Simulation of the wake of a fluidic pinball at a Reynolds number Re equal to 130, and PIV measurements of the wake of a $NACA\ 0018$ wing profile with $Re = 4\,800$. We show that the method is able to reconstruct the flow dynamics for horizons of several convective times, provided that the basis for the data reduction is sufficiently compact. Our results suggest that manifold learning and data assimilation can be combined to obtain an enhanced flow description..

1. Introduction

The advances of time-resolved (TR) Particle Image Velocimetry (PIV) have paved the way to a complete description of turbulent flows including pressure fields (Van Oudheusden, 2013). On the downside, TR PIV requires expensive equipment, and it is often affected by reduced laser light intensity and measurement noise due to the relatively larger pixel size of high-speed cameras, thus limiting its applicability. This has pushed towards developing methods to estimate pressure from non-time-resolved PIV. Methods based on Taylor's hypothesis (Laskari et al., 2016; Van der Kindere et al., 2019) or non-diffusive vorticity transport (Schneiders et al., 2016) demonstrated being capable of estimating pressure fields without time-resolution. In both cases a time-marching

procedure can be set to estimate pressure field sequences. The main issue is the progressive erosion of the domain due to advection through the boundaries. Furthermore, methods based on Taylor's hypothesis rely on a delicate choice of the correct convection velocity; non-diffusive vorticity transport, on the other hand, require 3D flow measurements to be implemented.

An alternative route to obtain a time-resolved flow description is to pour time resolution from point probes to non-time-resolved PIV measurements. Model-free approaches based on Extended Proper Orthogonal Decomposition (EPOD, Borée, 2003) have shown that the correlation between synchronized measurements can be leveraged in basis with reduced dimensionality for efficient estimation of one measured quantity as a function of another simultaneous measurement. Multi-time delay estimation can also be included to increase the robustness of the estimation (Tinney et al., 2008). This concept, with variants, has been exploited in wake flows (Bourgeois et al., 2013; Tu et al., 2013; Discetti et al., 2018), and turbulent wall-bounded flows (Hosseini et al., 2015; Discetti et al., 2019). Recently, Chen et al. (2022) have demostrated that this path can also be used to estimate pressure fields enforcing the Navier-Stokes equations. On the downside, this requires intrusive simultaneous measurements with PIV, which are often unfeasible due to access requirement, obstruction of the optical access, flow perturbation and unwanted reflections.

An interesting opportunity is offered by non-time-resolved multi-pulse PIV setups. Applications based on multiple synchronized illumination systems have been recently proposed (Lynch & Scarano, 2014; Ding & Adrian, 2016; Novara et al., 2019), either exploiting polarization to obtain single-exposed frames, or on multiple-exposed frames. This solution provides acceleration field measurements, although without time resolution.

In this work we investigate a possible pathway to obtain time-resolved measurement from non-time-resolved acceleration and velocity field measurements. The main principle we leverage is that turbulent flows show underlying patterns whose dynamics often evolve on low-dimensional manifolds. Provided that an adequate basis is identified, a parsimonious model for the flow dynamics can thus be sought to infer a time-marching estimation of the most significant flow scales. The method we propose is based on the sparse nonlinear system identification (SINDy) introduced by Brunton et al. (2016). Non-time-resolved instantaneous velocity and acceleration fields are fed into the system identification algorithm. First physical principles are then enforced to estimate the pressure field, in this case from spatial integration of pressure gradient from the Navier-Stokes equations. It must be remarked that this solution can also be fed by estimated fields from probe data, as in the work by (Loiseau et al., 2018). In this work we explore the possibility of using directly acceleration fields, without having to rely on probe measurements.

The proposed data-driven methodology is briefly described in §2. It will be shown that the method can be either used for short-term prediction, or to increase the temporal resolution by filling the temporal gaps between subsequent snapshots. In this last "time super-sampling" case, a backward-forward integration procedure is proposed. The potential of the technique is demonstrated in §3 on the DNS dataset of the wake of a fluidic pinball and on PIV experimental data of the wake past a stalled wing in a water tunnel.

2. Methodology

The flowchart of the proposed approach is sketched in Figure 1. It can be atomized in three main phases: dynamics estimation; time-marching reconstruction of the velocity fields; pressure computation.

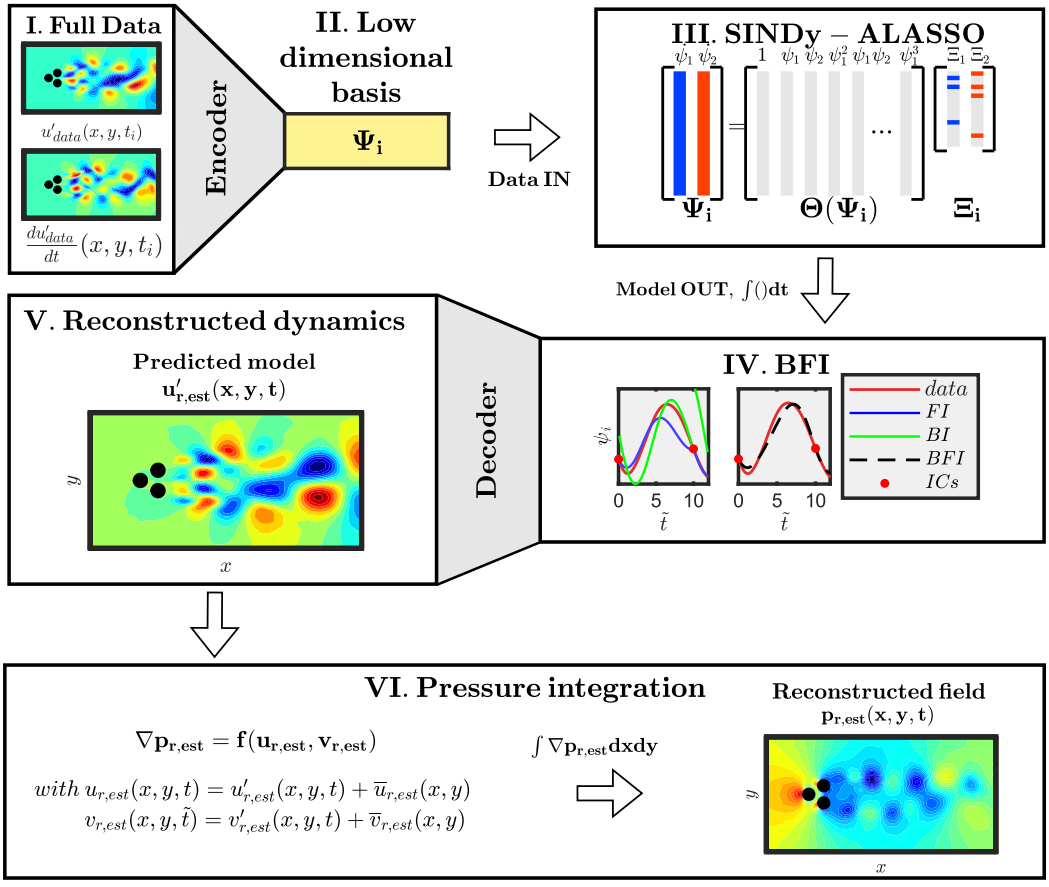


Figure 1. Schematic of the flow reconstruction algorithm. Non-time resolved data of velocity and acceleration fields are collected. Data are projected into a low-dimensional set of coordinates whose dynamics are inferred from a sparse system-identification method. Backward-forward integration is applied to the obtained model using the measured snapshots as initial conditions, providing an estimated time-resolved velocity field. Finally, the pressure field is obtained from the Navier-Stokes' equations.

2.1. Dynamics estimation

The dimensionality reduction of the velocity field data into a low-order basis is leveraged through Proper Orthogonal Decomposition (POD, Berkooz et al., 1993). POD seeks to find an optimal decomposition, in the sense of the minimization of the Frobenius norm of the residual energy of the decomposition. For discrete systems this can be achieved with Singular Value Decomposition (SVD) or the method of snapshots (Sirovich, 1987). Assume a snapshot matrix $\mathbf{U}[\mathbf{i},\mathbf{k}] \in \mathbb{R}^{N_s \times N_t}$, where $N_t \leq N_s$. The economy-size SVD arranges the data in a temporal history of N_t discretized

times such that $\mathbf{k} \in [1, ..., N_t]$ and in a spatial grid with $\mathbf{i} \in [1, ..., N_s]$, with N_s being the number of gridpoints times the vectorial dimension of the data.

$$\mathbf{U}[\mathbf{i}, \mathbf{k}] = \phi[\mathbf{i}]\mathbf{\Sigma}\mathbf{\Psi}[\mathbf{k}]^T. \tag{1}$$

In Equation (1), ϕ and $\mathbf{\Psi}$ refer to the spatial and temporal modes respectively, whereas $\mathbf{\Sigma}$ contains the singular values σ_j along its diagonal, sorted according to their value. The energy optimality of the POD renders it particularly appealing for the dimensionality reduction of the dynamics of velocity fields. POD temporal modes, for instance, have been shown to be suitable coordinates for parsimonious nonlinear dynamic estimation for shedding-dominated flows (Brunton et al., 2016) and shear-driven cavity flows (Callaham et al., 2022). Furthermore, Galerkin regression models can also be generalized to include constraints or higher-order effects of truncated terms (Loiseau & Brunton, 2018). A dynamical model will thus solely rely on seeking a formulation for a dynamical system as

$$\frac{d\mathbf{\Psi_r}}{dt}(t) = \mathbf{f}(\mathbf{\Psi_r}(t)) \quad with \quad \frac{d\mathbf{\Psi_r}}{dt} = \left(\mathbf{\Sigma_r}^{-1} {\phi_\mathbf{r}}^T \frac{d\mathbf{U}}{dt}\right)^T, \tag{2}$$

with vector $\mathbf{\Psi_r}(t) = [\psi_1(t) \dots \psi_{N_r}(t)] \in \mathbb{R}^{N_r}$ being the truncated temporal mode set at a time t. The time derivative of each mode is expressed as an unknown nonlinear function $f(\mathbf{\Psi_r}(t))$. The assumption of available velocity and acceleration measurements, although without temporal resolution, is necessary to infer the time derivative of the temporal modes $\frac{d\mathbf{\Psi_r}}{dt}(t)$ by projection of acceleration measurements on the POD basis, as shown in Equation (2).

The framework of Sparse Identification of Non-Linear Dynamics by Brunton et al. (2016) relies on the principle that many dynamical systems are sparse in the space of possible functions. SINDy identifies governing non-linear equations by performing a regularized linear regression model onto time series data, such that

$$\frac{d\mathbf{\Psi}_r}{dt} = \mathbf{\Theta}(\mathbf{\Psi_r})\mathbf{\Xi}. \tag{3}$$

Here, $\mathbf{\Theta} \in \mathbb{R}^{N_t \times N_f}$ is a library matrix of N_f candidate non-linear functions in terms of $\mathbf{\Psi_r}$, where typically $N_t \gg N_f$ defines the system as overdetermined. A dictionary based on polynomial functions has been shown to represent adequately turbulent flow behaviour (Loiseau et al., 2018), having a general form of

$$\mathbf{\Theta}(\mathbf{\Psi_r}) = [\mathbf{1} \quad \mathbf{\Psi_r} \quad \mathbf{P_2}(\mathbf{\Psi_r}) \quad \dots \quad \mathbf{P_n}(\mathbf{\Psi_r})], \tag{4}$$

where $\mathbf{P_n}(\mathbf{\Psi_r})$ refers to a matrix whose columns are given by all possible combinations of the states of the system that provide a n^{th} order polynomial. In Equation (3), each column of $\mathbf{\Xi} \in \mathbb{R}^{N_f \times N_r}$ denotes the active functions from the library basis for a given state.

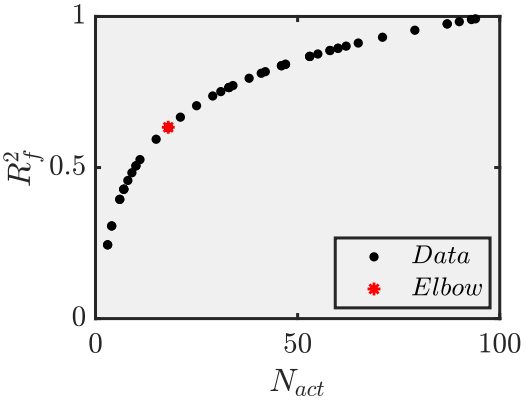


Figure 2. Expected behaviour of dynamical system fitting process. Each data point corresponds to a different value of the given parameter set, e.g. (δ, α) in ALASSO. The red point represents the best compromise between accuracy and sparsity according to the elbow criterion.

The active coefficients may vary depending on the regression model applied. Risk of overfitting the solution or encountering blow-up problems in the numerical integration of the dynamical system can be avoided by introducing appropriate regularization and sparsity-promoting parameters. Adaptive Least Absolute Shrinkage and Selection Operator (ALASSO, Zou, 2006), is reported by Fukami et al. (2021) to be a suitable and flexible candidate to enforce sparsity for high-dimensional complex flow problems.

$$\boldsymbol{\Xi}_{\mathbf{j}} = \underset{\boldsymbol{\Xi}'_{\mathbf{j}}}{argmin} \left|\left| \boldsymbol{\Theta}\boldsymbol{\Xi}'_{\mathbf{j}} - \frac{d\psi_{\mathbf{j}}}{dt} \right|\right|_2^2 + \alpha \sum_{j=1}^{N_r} (|\Xi'_j|)^{-\delta} |\Xi'_j|. \tag{5}$$

Equation (5) shows the optimization problem of the ALASSO algorithm. Note that it will always be sought to find the minimum amount of involved terms N_{act} in the dynamical system for the closest fit of Equation (2), that is, for a maximum determination coefficient R_f^2, where

$$R_f^2 = 1 - \frac{||z - z'||_2^2}{||z - \overline{z}'||_2^2}, \quad with \quad z = \frac{d\boldsymbol{\Psi}_r}{dt},\ z' = \boldsymbol{\Theta}(\boldsymbol{\Psi}_\mathbf{r})\boldsymbol{\Xi} \tag{6}$$

Regardless of the method, sparsity-enforced algorithms depend a set of free parameters (in the chosen case, $[\delta, \alpha]$) that have been proven (Brunton et al., 2016) to provide an intractable convergence to the desired optimal fit. Moreover, it must be noted that the minimum N_{act}/R_f^2 does not necessarily ensure a good retrieval of the dynamical features of the system. Indeed, the trend it presents has the shape of Figure 2. Therefore, selection of the mentioned parameters will be done through a brute-force iterative process that converges into a compromise between the two sought conditions $max(R_f^2)$ and $min(N_{act})$, corresponding to finding the elbow in Figure 2 (elbow criterion, Thorndike, 1953, Brunton et al., 2016).

2.2. Time-marching reconstruction of the velocity fields

From the retrieved dynamical system, a time-marching reconstruction of the velocity fields is obtained from Backward-Forward numerical integration (BFI) of Equation (2). Available non-time-resolved velocity field measurements are leveraged as a set of multiple Initial Conditions (ICs), thus obtaining a time super-sampling. BFI integrates the system forward ($\mathbf{\Psi}_{\mathbf{r},\mathbf{F}}(t)$) and backward ($\mathbf{\Psi}_{\mathbf{r},\mathbf{B}}(t)$) in time for each time span comprised between two adjacent PIV snapshots, that is, at times t_i and t_{i+1} where field measurements are available. A cubic B-spline function $\omega(t) \in [0,1]$ is introduced as a weighting factor that counteracts the propagation of each forward and backward integration error near t_{i+1} and t_i, respectively.

$$\begin{gathered}\psi_{j,est}(t) = (1-\omega(t)) \cdot \psi_{j,est,F}(t) + \omega(t) \cdot \psi_{j,est,B}(t) \quad with\, t \in [t_i, t_{i+1}] \\ \psi_{j,est,F}(t) = \int_{t_i}^{t} \mathbf{\Theta}(t)\mathbf{\Xi}_j \, dt \qquad \psi_{j,est,B}(t) = \int_{t}^{t_{i+1}} \mathbf{\Theta}(t)\mathbf{\Xi}_j \, dt\end{gathered} \tag{7}$$

2.3. Pressure estimation

The final step is attained by expressing the pressure gradient from the momentum equation in Equation (8) as a function of the estimated (indicated with subscript *est*) velocity fields $\nabla p = f(u_{est}, v_{est})$, obtained from BFI of its POD reduced-order dynamical model. The pressure field is integrated through the same iterative process described by Chen et al. (2022), based on a finite-differences version of the Modified Richardson Iteration method (Richardson, 1911).

$$\nabla p = -\rho \frac{D\mathbf{u}}{Dt} + \mu \nabla^2 \mathbf{u} \tag{8}$$

3. Validation

The proposed methodology will be validated through two different datasets: a numerical and an experimental one. For each flow configuration, its dynamical system will be inferred from a training set of snapshots. A different dataset will be then used as testing for the integration of the retrieved dynamical system.

3.1. DNS dataset: Fluidic Pinball

The synthetic test case is based on the Direct Numerical Simulation (DNS) of the wake of a fluidic pinball (Deng et al., 2020), comprised of three equally-spaced cylinders of normalized diameter D, whose centers form an equilateral triangle of side $3D/2$ (see Figure 3). All data will be hereon referred to their normalized values basen on the cylinder diameter D and the free-stream velocity

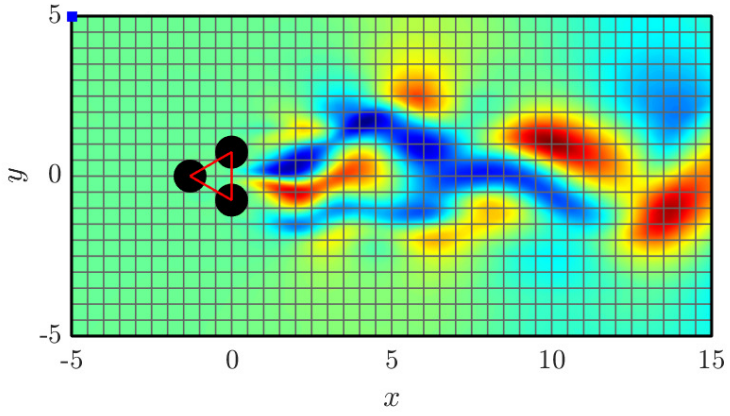


Figure 3. Generic snapshot displaying main geometrical features of the Fluidic Pinball dataset. A grid of 0.5 spacing in x and y is given for proportion purposes. The red equilateral triangle of side 1.5 shows the positioning of the three cylinders. The blue point represents the boundary condition used for the pressure integration.

Table 1. Main parameters of the dataset and the training case considered.

Dataset		Training	
$\Delta x, \Delta y$	0.1	N_t	1000
N_s	40602	N_r	10
Δt	0.1	$E\,[\%]$	90
Re	130	$n \in \mathbf{P_n}$	2

U. Table 1 provides a summary of the most relevant parameters of the discretized data and the training set.

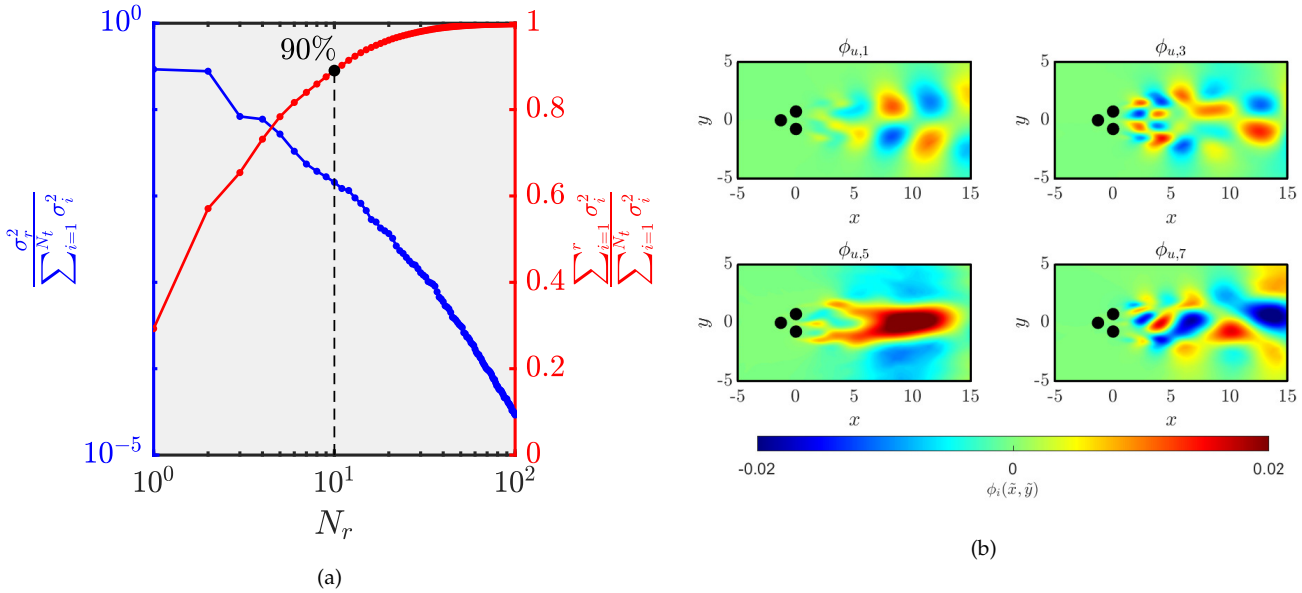
Velocity, acceleration and pressure data have been interpolated onto a regular grid ranging from $x = -5$, $y = -5$ to $x = 15$, $y = 5$ with a discretized spacing of $\Delta x = \Delta y = 0.1$. This spacing is considered sufficient for proper integration of the pressure gradient with minimal effect of truncation errors. The Reynolds number is set to $Re = 130$ and the time step adopted for the DNS is $\Delta t = 0.1$. This condition lies within the symmetric-chaotic regime (Deng et al., 2020). This regime is optimal for a first insight into chaotic behaviour estimation while still allowing for a representative enough low-order dimensionality reduction of the dynamics ensured by the dominance of the vortex shedding.

2D-velocity and acceleration data is considered for the estimation of the dynamical model in Equation (2). A training set of 1000 randomly-spaced snapshots with an average time spacing of $t = 1$ is prepared for this purpose. Reduction to a low-order model is chosen to be provided setting an energy threshold at approximately $90\ \%$, giving a system characterized by $r = 10$ modes. More refined methods based, for instance, on the elbow criterion (Thorndike, 1953), can be included to obtain a more robust user-independent rank selection. A parametric study of the optimal energy criteria based on reconstruction accuracy and low-order reduction is regarded for further develop-

ments.

For the selected flow regime the main flow features are an oscillating antisymmetric base-bleeding jet between the two rearward cylinders and the vortex shedding. The spatial modes in Figure 4b can be shown to separately distill the base flow contribution (mode 5) from the dissipation of vortices along the wake (modes 1,3,7).

A library basis of polynomials up to order 2 is constructed as a function of the truncated temporal mode set $\mathbf{\Psi_r}$. The given order has been chosen to be consistent with the basis spanned by Galerkin-POD projections (Noack et al., 2005). Although this second-order basis is proven to be complete only for a non-truncated model, it was assumed as the best compromise between a known analytical model and a reduction in computational cost in the training process. Expressing the dynamical model as Equation (3), the active functions for each temporal mode can be inferred from the ALASSO method in Equation (5). Notice in Figure 5 that the first 4 modes allow for a sparse representation of their dynamics, whereas the rest seem to be dominated by non-linearities. Moreover, it can be seen that modes 9 and 10 are not well dynamically represented with the available library matrix $\mathbf{\Theta}(\mathbf{\Psi_r})$.



(a) (b)

Figure 4. (a) Energy and cumulative energy distribution for the first 100 POD modes. **(b)** Contour maps of a selection of the most representative spatial POD modes of the horizontal velocity component.

Accuracy of the dynamic reconstruction is assessed through a testing dataset of two-dimensional velocity fields with varying time separation of $t_s \in [1, 10]$. Different degrees of undersampling between the instantaneous realizations are compared to provide an insight on the limitations of the numerical dynamic integration. A first estimation of the reconstruction error is thus done through a comparison of the relative root mean square error (RMSE, $\bar{\epsilon}$) between the estimated 2D-velocity fields and the *truncated DNS* for increasing time supersampling (see Figure 6a). The *truncated DNS* is here referred to a truncated version of the original DNS dataset to the same rank adopted for Equation (2). The original time-resolved DNS dataset will be hereafter denoted as *DNS*. The $\bar{\epsilon}$ for

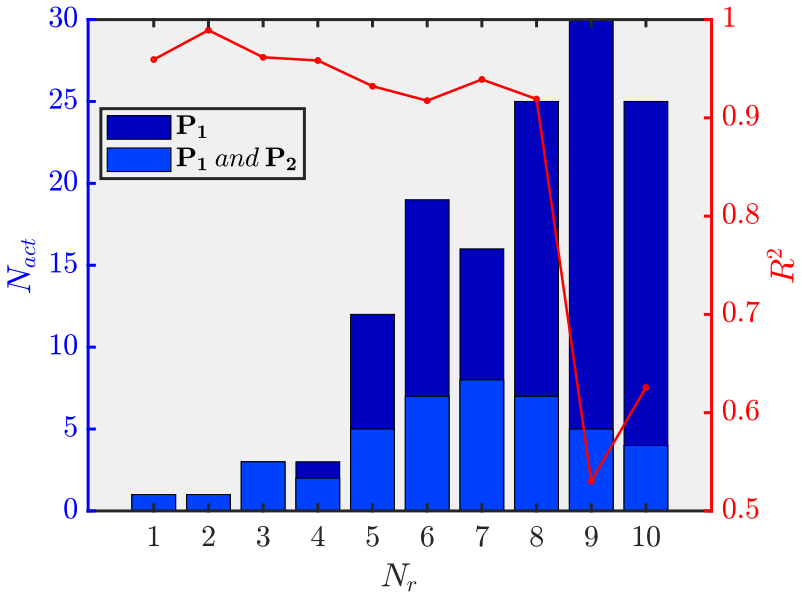


Figure 5. Results from ALASSO training. The bar plot represents the number of active terms of first ($\mathbf{P_1}$) and second order ($\mathbf{P_2}$). The right axis depicts the determination R^2 coefficient of the fitted dynamics.

any given snapshot matrix of data $\mathbf{U}[\mathbf{i},\mathbf{k}]$ is defined as

$$\bar{\epsilon}^2 = \frac{1}{N_s N_t \cdot \sigma_{ref}^2} ||\mathbf{U}[\mathbf{i},\mathbf{k}]_{est} - \mathbf{U}[\mathbf{i},\mathbf{k}]_{ref}||_F^2. \tag{9}$$

In Equation (9), the Frobenius norm $||\cdot||_F$ is computed onto the difference between the estimated and reference (subscript *ref*) data. The normalization parameter for the MSE computation is chosen as the variance σ_{ref}^2, evaluated on the reference data. Notice from Figure 6a that a weighted contribution of both forward and backward integration methods can result beneficial in the final determination of the velocity field evolution. In fact, BFI for a certain time spacing retrieves a more accurate description of the flow than what Forward Interpolation (FI) and Backward Interpolation (BI) are capable of doing for half the time separation. Except for reduced time separation, the proposed method shows superior performances if compared to cubic Spline Interpolation (SI).

The estimated velocity fields are then inserted in Equation (8) to estimate the pressure gradient, and thus, the pressure field with temporal resolution. Spatial integration of the pressure gradient requires a single-point boundary condition, provided by a added time-resolved history of a point pressure measurement, placed at $x = -5, y = 5$. The relative RMS error among the BFI estimated and *truncated DNS* pressure fields for increasing time separation of instantaneous realizations is included in Figure 6b. It can be shown that the pressure estimation error depends mostly on mode truncation, with the error due to inaccuracy of the low-rank velocity field estimated being relevant only for large time separation between measurements. It must be remarked that, even for significant time separations (up to $t_s = 5$), the error of the estimated pressure field is practically the same of performing the integration on the *truncated DNS* field. This suggests that the nonlinear system

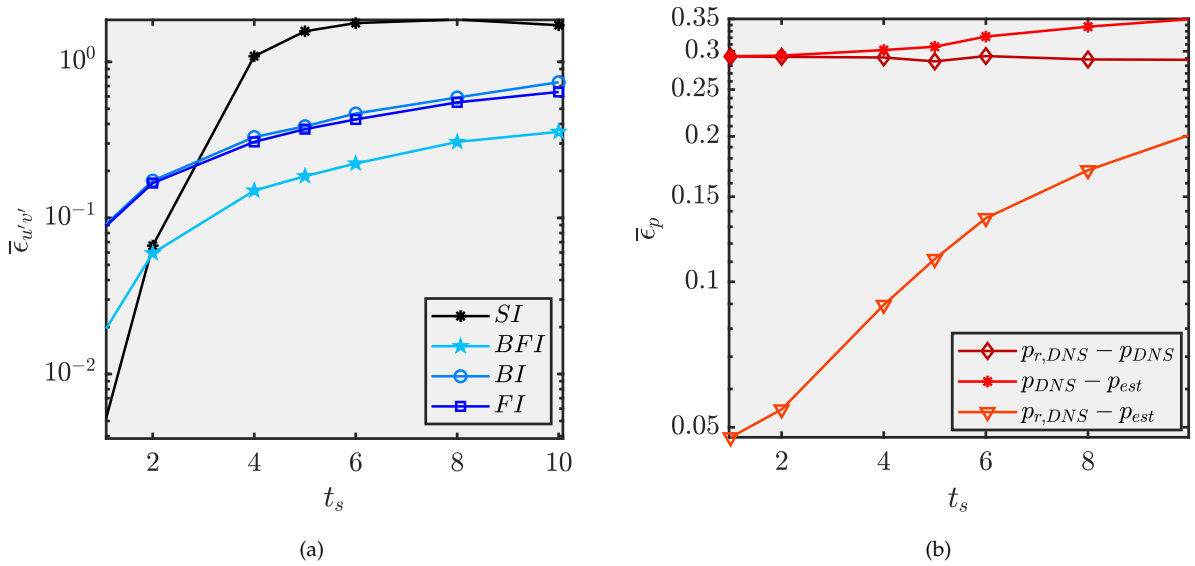


Figure 6. (a) Relative RMS error between the estimated and *truncated DNS* velocity fields for varying time spacing of the initial conditions. Cubic-Spline interpolation (SI), Forward (FI), Backward (BI) and Backward-Forward Integration (BFI) methods are compared. **(b)** Error of pressure field for several time separations between snapshots for estimated, *DNS* and *truncated DNS*. Velocity fields have been estimated with BFI method.

identification is capable of correctly identifying the dynamics generating the pressure fluctuations, within the limits of the performed rank truncation.

Further insight into the dynamical reconstruction capabilities is given with the results reported by a time separation case of $t_s = 5$, which corresponds to approximately one complete shedding cycle between consecutive measurements. A preliminary comparison of the backward-forward integration procedure with straightforward cubic spline interpolation of the POD time coefficients is illustrated in Figure 7. The quality of the fit is modelled through the squared correlation coefficient R^2 for each mode j (see Equation (10)). It is observed that, with the exception of some modes modeling slowly-evolving features (e.g. mode 5), the identified system is largely superior in accuracy with respect to POD coefficients interpolation.

$$R_j^2 = \frac{< \psi_{j,est}(t)\psi_{j,ref}(t) >^2}{< (\psi_{j,est}(t)^2 >< (\psi_{j,ref}(t))^2 >}. \tag{10}$$

Figure 8 depicts a reconstructed pressure field computed at a time instant equidistant between two consecutive realizations, i.e. at $t = 2.5$, for which the estimation error on the integrated dynamical model is expected to be maximum. It can be seen that the pressure integration based on the estimated BFI velocity fields largely succeeds in identifying the most relevant features of the pressure distribution.

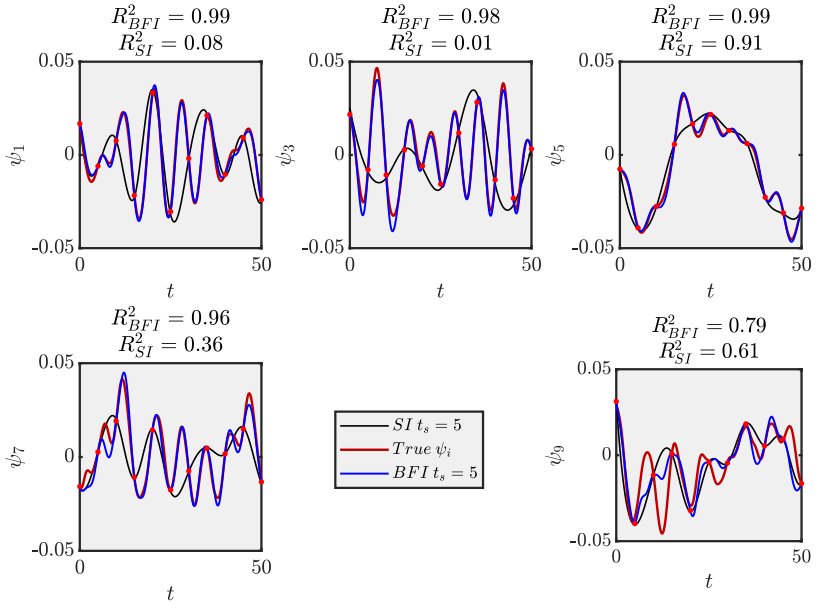


Figure 7. Reconstructed POD temporal modes from Backward-Forward Integration method with snapshot separation of $t_s = 5$. The time at which the initial conditions are taken are depicted as red dots. The correlation coefficient R^2 is compared between the backward-forward integration and a cubic spline interpolation.

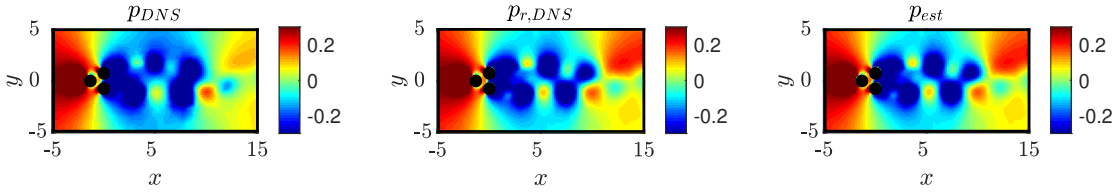


Figure 8. Contour of the pressure distributions from integration of Equation (8) using velocity and acceleration from the *DNS* (left), *truncated DNS* (middle) and estimated field with BFI (right). A case with time separation $t_s = 5$ is considered. The pressure field corresponds to a snapshot separated by $t = t_s/2$ from measured ones, i.e. equidistant from two realizations.

3.2. Experimental dataset: wake on wing profile

An experimental validation is performed using planar PIV measurements on the wake of a 2D stalled wing with a NACA 0018 airfoil. The experimental setup is briefly described here; full details are available in the work by Chen et al. (2022). The experiments are carried out in the water tunnel of the Universidad Carlos III de Madrid. A wing chord length $L = 80\,mm$ and a free-stream velocity of $U_\infty = 6\ cm/s$ are set, thus resulting in a Reynolds number of 4800. Spatial resolution is set to $R_s = 0.1197\ mm/px$ with a field of view covering $x \in [0.75, 2.4]$ and $y \in [-0.75, 0.75]$ (with x and y being non-dimensional coordinates normalized with the wing chord L). Figure 9 and Table 2 provide for the main characteristics of the exploited dataset.

The original sampling frequency of $f_s = 30\ Hz$ (i.e. $\Delta t = 0.025$) provides a time-resolved description of the flow field. In effect, the original PIV resolution will be leveraged as the reference dataset

(hereon referred as PIV) for comparison with the reconstructed flow from non-time resolved measurements obtained by downsampling the original sequence. The training set is composed of 1000 snapshots with an average time spacing of $t_s = 0.375$.

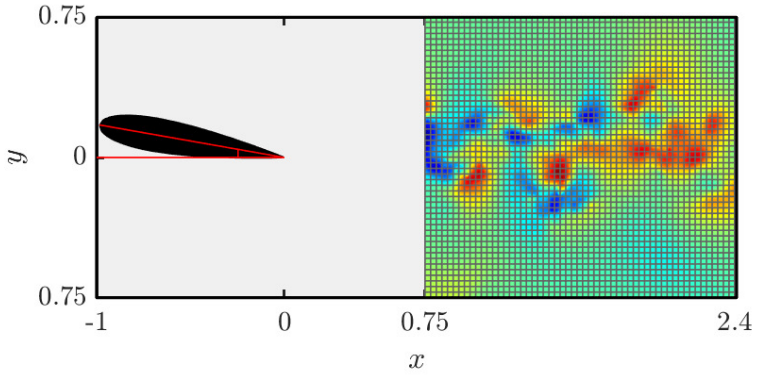


Figure 9. Generic snapshot displaying main geometrical features of the wing experiment dataset. The gridded region corresponds to the span of the available 2D data. A 0.06 spacing in x and y is given for proportion purposes. The red lines indicate the orientation of the chordline of the $NACA\ 0018$ airfoil with an angle of attack of $\alpha = 10°$.

Table 2. Main parameters of the dataset and the training case considered.

Dataset				Training	
Dimensional		Normalized			
$L\ [mm]$	80	$\Delta x, \Delta y$	0.015	N_t	1000
$U_\infty\ [cm/s]$	6	N_s	22624	N_r	41
$R_s\ [mm/px]$	0.1197	Δt	0.025	$E\ [\%]$	80
$f_s\ [Hz]$	30	Re	4800	$n \in \mathbf{P_n}$	1

Differently from the previous test case, a less compact basis transformation is now retrieved; only about 80% of the energy is captured by 41 POD modes, thus entailing initial constraints in terms of complexity of the dynamical system. The library basis for the system identification process is limited now to first-order polynomials to reduce the computational cost. Higher-order basis were shown to involve a significant increase in the computational effort of the available optimization procedure, while not providing a significant improvement in terms of accuracy of the system identification. Figure 10a shows the number of active terms and the determination coefficient of each temporal mode dynamics. It can be seen that efficiency in achieving a simple sparse system determination of a turbulent case substantially decreases with respect to the previous synthetic test case.

However, combined integration of the retrieved model with BFI can be shown to provide a similar reconstructed flow field to the reference time-resolved sequence. The dynamical model is integrated using as initial conditions data from a testing dataset of two-dimensional velocity fields

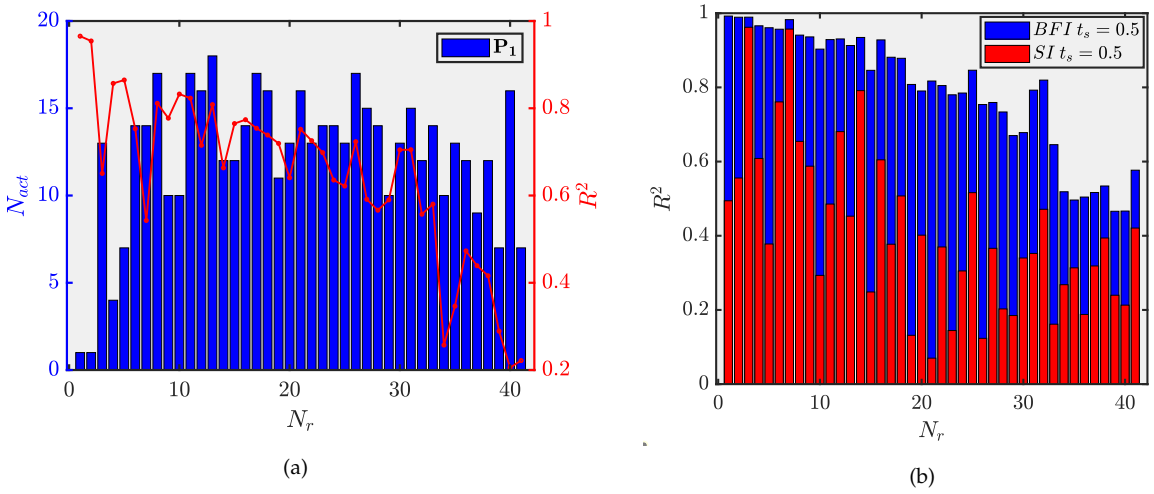


Figure 10. (a) Results from ALASSO training for the wing training dataset. **(b)** Correlation coefficient R^2 of the reconstructed POD temporal modes using BFI and SI, for a time super-sampling of $t_s = 0.5$.

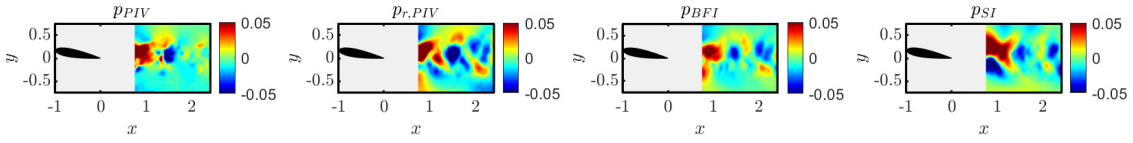
with two larger degrees of undersampling: $t_s = [0.5, 1]$, corresponding to super-sampling factors of $\mathcal{F}_s = [20, 40]$.

BFI and SI of the POD temporal basis is contrasted in Figure 10b for one of the time separation cases. The identified dynamics for the first modes succeeds in providing an accurate estimation of its time history. Higher-order modes are not captured as sharply, although still BFI shows being superior to SI.

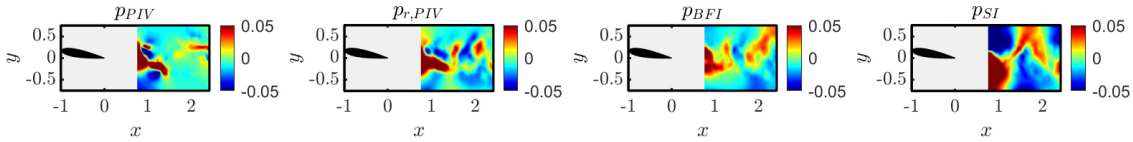
The reconstructed instantaneous pressure field is graphically compared in Figure 11 with the reference time-resolved sequence and the truncated version of the latter (*truncated PIV*) for the same energy threshold used in estimation. The instantaneous fields are chosen at the most critical time instant in terms of integration error, that is, at the equidistant point between consecutive realizations. Time-marching pressure representation of the BFI estimated field proves to capture the most relevant dynamical features of the system in the limit of its energy truncation capabilities, and it shows remarkably higher performances than SI.

4. Conclusions

An enhanced flow description method from incomplete measurements has been presented. The proposed technique roots its basis on nonlinear system identification and on the availability of velocity and acceleration field measurements with poor time-resolution. We demonstrated the time propagation being feasible with a model-free approach, i.e. the dynamics is directly extracted from the data. Enforcing first physical principles is then possible to compute time-resolved pressure fields, thus disclosing relevant information, for instance, for unsteady aerodynamics or aeroacoustic noise investigation. The time history of the velocity and pressure fields is obtained through its dynamic estimation and a weighted integration process between consecutive non-time-resolved



(a)



(b)

Figure 11. From left to right, contour of pressure distribution of the reference time-resolved sequence, the *truncated PIV*, the estimated field with BFI and with SI. The pressure field corresponds to a snapshot separated by $t = t_s/2$ from measured ones, i.e. equidistant from two realizations. Two cases of time separations of **(a)** $t_s = 0.5$ and **(b)** $t_s = 1$ are considered.

realizations. The algorithm is validated on 2D simulation data of the wake of a fluidic pinball at $Re = 130$. Notwithstanding its simplicity, this test case features a chaotic behaviour, thus posing already a challenge to the identification of sparse dynamical systems. Extension to more complex turbulent flows is attained with a two-dimensional experimental dataset of the wake of a wing profile with $Re = 4800$. The increase in complexity in its POD basis is a challenge for the convergence of the system identification technique. Nevertheless, for both scenarios, results are shown to capture the most relevant dynamics for significant time separation values and excel those from regular spline interpolation of POD time coefficients between non-time-resolved snapshots.

Acknowledgements

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 949085). The authors warmly acknowledge M. Morzynski and B. Noack for granting access to the fluidic pinball DNS code, and M. Raiola and J. Chen for providing the 2D wing experimental data and the pressure integration code.

References

Berkooz, G., Holmes, P., & Lumley, J. L. (1993). The proper orthogonal decomposition in the analysis of turbulent flows. *Annual review of fluid mechanics*, *25*(1), 539–575.

Borée, J. (2003). Extended proper orthogonal decomposition: a tool to analyse correlated events in turbulent flows. *Experiments in Fluids*, *35*(2), 188–192.

Bourgeois, J., Noack, B., & Martinuzzi, R. (2013). Generalized phase average with applications to sensor-based flow estimation of the wall-mounted square cylinder wake. *Journal of Fluid Mechanics*, *736*, 316–350.

Brunton, S. L., Proctor, J. L., & Kutz, J. N. (2016). Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the national academy of sciences*, *113*(15), 3932–3937.

Callaham, J. L., Brunton, S. L., & Loiseau, J.-C. (2022). On the role of nonlinear correlations in reduced-order modelling. *Journal of Fluid Mechanics*, *938*.

Chen, J., Raiola, M., & Discetti, S. (2022). Pressure from data-driven estimation of velocity fields using snapshot PIV and fast probes. *Experimental Thermal and Fluid Science*, *136*, 110647.

Deng, N., Noack, B. R., Morzyński, M., & Pastur, L. R. (2020). Low-order model for successive bifurcations of the fluidic pinball. *Journal of fluid mechanics*, *884*.

Ding, L., & Adrian, R. J. (2016). N-pulse particle image velocimetry-accelerometry for unsteady flow-structure interaction. *Measurement Science and Technology*, *28*(1), 014001.

Discetti, S., Bellani, G., Örlü, R., Serpieri, J., Vila, C. S., Raiola, M., … Ianiro, A. (2019). Characterization of very-large-scale motions in high-Re pipe flows. *Experimental Thermal and Fluid Science*, *104*, 1–8.

Discetti, S., Raiola, M., & Ianiro, A. (2018). Estimation of time-resolved turbulent fields through correlation of non-time-resolved field measurements and time-resolved point measurements. *Experimental Thermal and Fluid Science*, *93*, 119–130.

Fukami, K., Murata, T., Zhang, K., & Fukagata, K. (2021). Sparse identification of nonlinear dynamics with low-dimensionalized flow representations. *Journal of Fluid Mechanics*, *926*.

Hosseini, Z., Martinuzzi, R. J., & Noack, B. R. (2015). Sensor-based estimation of the velocity in the wake of a low-aspect-ratio pyramid. *Experiments in Fluids*, *56*(1), 13.

Laskari, A., de Kat, R., & Ganapathisubramani, B. (2016). Full-field pressure from snapshot and time-resolved volumetric PIV. *Experiments in fluids*, *57*(3), 1–14.

Loiseau, J.-C., & Brunton, S. L. (2018). Constrained sparse Galerkin regression. *Journal of Fluid Mechanics*, *838*, 42–67.

Loiseau, J.-C., Noack, B. R., & Brunton, S. L. (2018). Sparse reduced-order modelling: sensor-based dynamics to full-state estimation. *Journal of Fluid Mechanics*, *844*, 459–490.

Lynch, K., & Scarano, F. (2014). Material acceleration estimation by four-pulse tomo-PIV. *Measurement Science and Technology*, *25*(8), 084005.

Noack, B. R., Papas, P., & Monkewitz, P. A. (2005). The need for a pressure-term representation in empirical Galerkin models of incompressible shear flows. *Journal of Fluid Mechanics*, *523*, 339–365.

Novara, M., Schanz, D., Geisler, R., Gesemann, S., Voss, C., & Schröder, A. (2019). Multi-exposed recordings for 3D lagrangian particle tracking with multi-pulse shake-the-box. *Experiments in Fluids*, *60*(3), 1–19.

Richardson, L. F. (1911). The approximate arithmetical solution by finite differences with an application to stresses in masonry dams. *Philosophical Transactions of the Royal Society of America*, *210*, 307–357.

Schneiders, J. F., Pröbsting, S., Dwight, R. P., van Oudheusden, B. W., & Scarano, F. (2016). Pressure estimation from single-snapshot tomographic PIV in a turbulent boundary layer. *Experiments in Fluids*, *57*(4), 1–14.

Sirovich, L. (1987). Turbulence and the dynamics of coherent structures. I. Coherent structures. *Quarterly of applied mathematics*, *45*(3), 561–571.

Thorndike, R. L. (1953). Who belongs in the family. In *Psychometrika.*

Tinney, C., Ukeiley, L., & Glauser, M. N. (2008). Low-dimensional characteristics of a transonic jet. Part 2. Estimate and far-field prediction. *Journal of Fluid Mechanics*, *615*, 53–92.

Tu, J. H., Griffin, J., Hart, A., Rowley, C. W., Cattafesta, L. N., & Ukeiley, L. S. (2013). Integration of non-time-resolved PIV and time-resolved velocity point sensors for dynamic estimation of velocity fields. *Experiments in Fluids*, *54*(2), 1–20.

Van der Kindere, J., Laskari, A., Ganapathisubramani, B., & De Kat, R. (2019). Pressure from 2D snapshot PIV. *Experiments in fluids*, *60*(2), 1–18.

Van Oudheusden, B. (2013). PIV-based pressure measurement. *Measurement Science and Technology*, *24*(3), 032001.

Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American statistical association*, *101*(476), 1418–1429.