An efficient offline sensor placement method for flow estimation

Junwei Chen^{a,*}, Marco Raiola^a, Stefano Discetti^a

^aDepartment of Aerospace Engineering, Universidad Carlos III de Madrid, Avda. Universidad 30, 28911, Leganes, Spain

Abstract

We present an efficient method to optimize sensor placement for flow estimation using sensors with time-delay embedding in advection-dominated flows. Our solution allows identifying promising candidates for sensor positions using solely preliminary flow field measurements with non-time-resolved Particle Image Velocimetry (PIV), without introducing physical probes in the Now Data-driven estimation in advection-dominated flows often exploits time-delay embedding to enren the sensor information. Optimizing for the reconstruction, i.e. it uses the information embedded in probe time series to provide a more accurate estimation. Optimizing the probe position is the key to improving the accuracy of such estimation. Unfortunately, the cost of performing an online combinatorial search to identify the optimal sensor placement in experiments is often prohibitive. We leverage the principle that, in advection-dominated flows, rows of vectors from PIV fields embed similar information to that of probe time series located at the domstream end of the domain. We propose thus to optimize the sensor placement using the row data from non-time-resolved field estimation. We show that the probesc and be installed outperforms equidistant positioning or greedy optimization techniques available in the literature.
Keywords: Particle Image Velocimetry, Proper Orthogonal Decomposition, Data-driven estimation
Introduction Particle Image Velocimetry (PIV) is referred to as "time-resolved" when the repetition rate of illumination and imaging system is fast enough to enable measurements of the flow dynamics with sufficient temporal resolution. Time-resolved flow dynamics (Geresh, 2021), unlocking a detailed view of flow dynamics (Geresh, 2021), unlocking a detailed view of flow dynamics. Sufficiently fast.
As a cost-effective and Rexibe alternative, high-repetition rate point probes (such as hot wires and microphones) can be everaged to set and probaches leverage the formelest as a dimensolution. The work provided interesting guidelines and proceduced with simultaneous low-repetition-rate PIV measurements. Data-driven techniques can then be leveraged to estimation. The work provided interesting guidelines and proceduced with sensors. The aforementioned approaches leverage the principrobe data. The reconstruction problem is ill-posed: few sens flow. Data-driven estimation in advection-dominated flows often exploits time-delay embedding to enrich the sensor information of the reconstruction, i.e. it uses the information embedded in probe time series to provide a more accurate estimation. Optimizing

sors are used to estimate flow fields composed of a much larger number of vectors. However, it can still be resolved with reasonable accuracy according to the following two observations. First, often the flow dynamics evolve on low-dimensional attractors. Dimensionality-reduction techniques, such as Proper

sors. Additionally, the estimated time-resolved flow fields can be exploited to retrieve additional field information. Recently, it was demonstrated that pressure fields can also be obtained from the estimated velocity fields by leveraging compliance with the Navier-Stokes equations and setting proper boundary conditions (Chen et al., 2022; Moreno Soto et al., 2024).

The sensor positioning has a crucial impact on the quality of the flow estimation. Proper placement improves the accuracy of flow field predictions and/or allows reducing the number

^{*}Corresponding author

Email addresses: junwei.chen@uc3m.es (Junwei Chen),

mraiola@ing.uc3m.es (Marco Raiola), sdiscett@ing.uc3m.es (Stefano Discetti)

of required sensors, thus simplifying the measurement setup. The main difficulty is that sensor placement requires solving a NP-hard (non-deterministic polynomial-time hard) combinatorial problem. Several methods have been proposed to solve it with an affordable computational cost. Remarkable examples include: QR-pivoting (Manohar et al., 2018), building upon the discrete empirical interpolation method (Drmac and Gugercin, 2016; Chaturantabut and Sorensen, 2010), Shannon's entropy (Papadopoulou et al., 2014), variance inflation factor (Nickels et al., 2021) and reinforcement learning (Paris et al., 2021). The main issue is that these methods do not account for the intrusiveness of the sensor. This is a key aspect, which most often leads to sensor placement at walls, or at the downstream end of the domain if advection is dominant.

In this work, we treat optimal sensor placement in this latter case, in which there is a main advection direction and point probes can be placed right at the downstream end of the domain. This introduces further complexity to the sensor optimization problem. The reconstruction accuracy, in this scenario, depends indeed also on the effectiveness of time-delay embedding in the selected locations. Such analysis can only be done a posteriori (i.e. after probe positioning and data acquisition). This would require performing the experiment multiple times with different probe positioning and searching for the best location. Even with a small handful of probes, this implementation is unaffordable.

In this paper, we propose an offline solution for sensor positioning, based solely on non-time-resolved PIV fields, i.e. without placing physical probes in the first preparatory run. Synthetic point probes are extracted from PIV fields to simulate the effect of including time history when placing the real ones. Using data from the preparatory PIV run, the flow field is reconstructed with different combinations of sensor locations on the downstream edge of the domain to identify the optimal positioning with the highest reconstruction accuracy. Once the most performing sensor placement is identified, the estimation experiment can be carried out by placing the physical probes and measuring the velocity fields simultaneously, and time-resolved flow field can be reconstructed from the measurement using PIV and sensors together.

The paper is structured as follows. First, the methodology is described in Sec. 2, including the estimation of EPOD, the process of online and offline sensor placement optimization, sensor formations, as well as block-pivoted QR for sensor placement. Then the validation test cases are illustrated (Sec. 3) and the results are discussed (Sec. 4). Finally, the conclusions are drawn.

2. Methods

2.1. Flow field estimator based on Extended Proper Orthogonal Decomposition

For the flow field estimation from probes, we use EPOD, a fast and reliable method that allows efficient exploration of sensor placements. While this paper focuses on EPOD, the same approach can extend to other estimators, e.g. those based on recurrent neural networks (Deng et al., 2019).

The EPOD (Boree, 2003) is a linear tool to establish the correlation between different quantities. In this framework, it is used to project the velocity field data from snapshot PIV on POD modes of the high-repetition-rate probe data at the same time instants. This allows the establishment of correlations between temporal modes of velocity field and probe data. The correlation can be leveraged then to estimate time coefficients of the velocity fields solely from probe data, thus delivering flow estimation.

This method is computationally affordable and provides acceptable estimation in simple flows (i.e. with a sufficiently compact POD spectrum and good evidence of correlation between sensor and field data, such as in shedding-dominated flows). The field estimation methodology is briefly described in this section. The reader can refer to Discetti et al. (2018) for a more detailed formulation. Furthermore, neural networks could also be used for this task (Erichson et al., 2020), although at the expense of a higher computational cost.

Consider a dataset of n_t snapshots of the fluctuating velocity field. The velocity is measured on n_p grid points. A snapshot matrix **U** is built by arranging each snapshot as a row. Following the snapshot method from Sirovich (1987), the POD is obtained by decomposing U with an economy-size singular value decomposition (SVD), i.e.

$$\mathbf{U} = \mathbf{\Psi} \mathbf{\Sigma} \mathbf{\Phi}^T \tag{1}$$

It is assumed for simplicity $n_t < n_c n_p$ (being n_c the number of velocity components measured at each grid point). The decomposition of Eq. 1 generates two square unitary matrices: Ψ , of size $n_t \times n_t$, containing the temporal modes; Φ^T containing the spatial modes, of size $n_c n_p \times n_c n_p$ (being it actually of size $n_t \times n_c n_p$ in the economy-size SVD for the case of $n_t < n_c n_p$). The diagonal matrix Σ contains the singular values σ_i sorted by their magnitude.

We apply the same decomposition to datasets acquired at the same time instants of velocity fields and probe measurements. As mentioned above, the number of probes is artificially increased by a multi-time-delay embedding. This approach is also referred to as Multichannel Singular Spectrum Analysis (Ghil et al., 2002). In particular, being the probes generally located at the downstream end of the domain, for each physical probe and at each time instant we assign a portion of the timeresolved sequence of the probe of q samples corresponding to time instants past the one of each corresponding PIV snapshot. This multi-time-delay embedding results in a probe snapshot matrix \mathbf{U}_{pr} with n_t rows and $n_{tt} = (s \times n_{c_{pr}}) \times q$ columns (being s the number of probes and $n_{c_{pr}}$ the number of flow quantity components measured by the probes). This matrix undergoes the same decomposition of the velocity snapshot matrix, resulting in:

$$\mathbf{U}_{pr} = \mathbf{\Psi}_{pr} \mathbf{\Sigma}_{pr} \mathbf{\Phi}_{pr}^{T} \tag{2}$$

The correlation matrix of temporal modes is thus built as $\Xi = \Psi_{pr}^T \Psi$. The matrix Ξ establishes the correlation between the velocity field and probe temporal modes.

The estimation of the velocity field at a generic time instant t^* is based on three steps: (1) computation by projection of the

temporal modes of the probes corresponding to t^* ; (2) estimation of the temporal modes of the velocity fields ψ_{est} leveraging their relation with the temporal modes of the probes through the correlation matrix Ξ ; (3) calculation of the velocity field using the estimated temporal modes and the spatial modes from the original decomposition. It can be shown that this can be condensed in the following relation:

$$\mathbf{u}_{est} = \mathbf{u}_{se} \mathbf{\Phi}_{pr} \boldsymbol{\Sigma}_{pr}^{-1} \boldsymbol{\Xi} \boldsymbol{\Sigma} \mathbf{\Phi}^{T} = \boldsymbol{\psi}_{est} \boldsymbol{\Sigma} \mathbf{\Phi}^{T}$$
(3)

with \mathbf{u}_{se} being a vector containing delay-embedded probe data arranged with the same procedure used to build the probe snapshot matrix. Further improvement of the estimation can be obtained by filtering out spurious contributions due to weakly-correlated features, following the method proposed by Discetti et al. (2018).

In this paper, the EPOD will be used for both training and testing. During the training phase, it will establish the relationship between the velocity field and sensor data using non-timeresolved data. In the testing phase, the estimated time-resolved field will be compared to the reference field.

2.2. Online vs. offline sensor placement optimization

Data-driven methods, such as the EPOD described in the previous section, can be used to provide temporal resolution to snapshot PIV if coupled with high-repetition-rate sensors. An accurate flow estimation allows velocity fields to be obtained ideally at the same temporal resolution as the probes. A sound placement of the sensors is paramount to ensure the quality of the temporal reconstruction of the flow field. Flow estimation includes two phases: i) sensor placement (either based on preliminary tests or on intuition of the user); ii) final testing using both PIV and sensors simultaneously to estimate the timeresolved flow field.

The most straightforward method for optimizing sensor placement is to deploy real probes during the first phase and evaluate their performance, a process we refer to as **"online optimization"**. This approach involves the following steps:

- 1. placing the probes in the candidate location;
- capturing and processing the data from PIV and from the probes;
- 3. performing the estimation process;
- 4. evaluating the quality of estimation and refining the sensor placement until an optimum in the accuracy of the estimation is reached.

It must be remarked that, even with a small number of probes and limited candidate locations, this approach becomes quickly unaffordable from the practical viewpoint. For *p* probes and *n* candidate locations, the number of required PIV experiments is n!/((n - p)!p!) for a full coverage of all combinations.

To address this, we propose an "offline optimization" method for sensor placement, relying solely on one dataset of non-time-resolved PIV data for the sensor placement, i.e. without requiring physical probes. This approach is sketched in Fig. 1. A similar principle is leveraged in other sensor optimization frameworks based on pattern discovery (for instance

in Manohar et al., 2018). However, a challenge arises here because delay embedding requires sensor signals or pointwise velocity from time-resolved PIV, which snapshot PIV cannot provide. The proposed method obviates the need for temporal resolution by using surrogates of the sensor time series extracted from snapshot PIV. In the following section, the nature of these surrogates as well as their difference with real probes time-series will be clarified.

2.3. Definition of sensor formations

In advection-dominated flows, sensors are typically placed near the downstream boundary of the domain in order to minimize the (1) wake produced inside the domain, (2) the reflection on the sensors in PIV images. In certain applications, placing sensor supports downstream may not be practical, leading to sensors being installed beneath the surface of objects. However, this paper does not consider that specific scenario. As mentioned in Sec. 2.1, the probe number can be artificially increased by using time segments of signal. The underlying hypothesis is that, if advection is dominating, the velocity recorded by the probe in time will be similar to the velocity of the flow upstream at earlier time instants. In this paper, this approach, used by among others Hosseini et al. (2015); Discetti et al. (2018) will be referred to as "probes time series". This arrangement is possible only by deploying physically the probes in the domain and performing synchronized measurements with PIV.

To avoid the time-consuming process of performing multiple probe deployments to identify the optimum sensor placement, we introduce a proxy for physical probe time series by exploiting the temporal and spatial coherence of advection-dominated flows. Specifically, we simulate probe time series using data from an entire row in PIV snapshots. These simulated sensors, referred to as "**row surrogate sensors**", are based on the assumption that the flow advects predominantly in the horizontal direction without loss of generality. It is important to note that the row surrogate sensors are usually extracted from single snapshots of non-time-resolved PIV fields, thus they are only useful for offline optimization and cannot enhance the temporal resolution of flow fields.

A potential limitation of the row surrogate sensors is that Taylor's hypothesis should be valid throughout at least one convection through-flow time (i.e. the time it takes to the flow to cross the measurement domain). However, it can be expected that this hypothesis generally holds near the probe location, and then becomes less reliable farther upstream. To mitigate this, we propose masking the row sensors based on their correlation with the downstream points of the row (candidate position for the probes). The correlation map $R_{ij,iN}$ between the generic point and the last position in the same row is given by,

$$R_{ij,iN} = \frac{\langle u'_{ij}u'_{iN} \rangle}{\sqrt{\langle u'^2_{ij} \rangle \langle u'^2_{iN} \rangle}}$$
(4)

where u'_{ij} is the streamwise fluctuation velocity at the *j*-th position of the *i*-th row and u'_{iN} is the streamwise fluctuation

ONLINE OPTIMIZATION



Figure 1: The workflow of online (top) and offline (bottom) sensor placement optimization.

velocity at the downstream end of the same row. The operator $\langle \cdot \rangle$ indicates the temporal average. For planar PIV, this equation produces a two-dimensional (2D) distribution, while for volumetric PIV, it can be calculated for each cross-section, or averaged across the volume thickness to improve convergence. To mask row sensors, we retain only the points with correlation values above a threshold, ensuring that at least 75% of rows retain more than 75% of their entries. These masked sensors, termed **"masked row sensors"**, are expected to better simulate probe time series by excluding less correlated regions.

To summarize, three kinds of sensor formation will be considered:

- **Probes time series:** these are the time-series extracted from physical probe measurements (i.e. from probes placed in the flow) as described in Sec 2.1. These data would be the ones available during an online optimization. It is the most intensive from the position optimization viewpoint as testing each combination requires a complete experiment.
- Row surrogate sensors: these are the surrogate timeseries used by our proposed "offline" probe position optimization. It is directly applied to the snapshot PIV data by extracting rows of vectors and using them to reconstruct the full field. The combination with the highest reconstruction quality defines the optimal probe location. This method requires 2 experiments: (1) a standard PIV experiment for offline optimization, without any probe deployment; (2) an experiment with synchronized PIV and probes in the optimal location identified in (1) to perform the time-delay embedding and corresponding flow estimation of time-resolved fields.
- Masked row sensors: these surrogate time-series are analogous to the ones used in offline optimization with

row surrogate sensors, the main difference being a masking applied to exclude points with low correlation from the process. The main hypothesis is that points with high correlation within each row would better represent the actual recording of physical probes capturing time series.

2.4. A quick sensor placement method

The proposed offline sensor placement optimization is based on a brute-force search of possible combinations of rows to obtain the best candidates for flow estimation once the probes are located in the field. This approach is practical if two conditions are met: (1) the estimation process is computationally efficient (e.g., the EPOD method used here); (2) the number of candidate locations and sensors is limited. For example, with 20 candidate positions and a set of 3 probes, the number of possible combinations exceeds 1000, which remains computationally manageable.

There are already some method to determine the sensors' position without such brute-force search. We explore an alternative greedy optimization method based on a modified version of QR-pivoting, as described in Manohar et al. (2018). This method is derived from the QR discrete empirical interpolation method (Drmac and Gugercin, 2016).

QR-pivoting is briefly explained here; readers can refer to Manohar et al. (2018) for more details.

Consider a full-field data snapshot represented as a vector $\mathbf{u} \in \mathbb{R}^{1 \times n_p}$. For simplicity, we assume only one measured component, though this can be extended to multiple components. If *s* sensors are available, the sensor measurements $\mathbf{p} \in \mathbb{R}^{s}$ can be expressed as,

$$\mathbf{p} = \mathbf{C}\mathbf{u}^{\mathrm{T}} \tag{5}$$

Here, the measurement matrix $\mathbf{C} \in \mathbb{R}^{s \times n_p}$ consists of zeros and ones, representing whether a sensor is located at that grid point

of the velocity field. The field **u** can also be expressed as a linear combination of a basis $\mathbf{\Phi} \in \mathbb{R}^{n_p \times n_t}$,

$$\mathbf{u}^T = \mathbf{\Phi} \mathbf{a}^T, \qquad \mathbf{a} \in \mathbb{R}^{1 \times n_t} \tag{6}$$

where **a** is a vector of coefficients. As in the previous section, we assume $n_t < n_p$. The basis $\mathbf{\Phi}$ can be predefined (e.g., Fourier modes) or data-driven (e.g., derived from POD). If using POD, **a** contains the time coefficients in any snapshot weighted by singular values. Substituting this into the sensor measurement equation gives

$$\mathbf{p} = (\mathbf{C}\boldsymbol{\Phi})\mathbf{a}^T \tag{7}$$

If the basis includes a number of modes larger than the number of sensors, the system becomes underdetermined. To address this, we reduce the basis to the first *r* modes, $\mathbf{\Phi}_r \in \mathbb{R}^{n_p \times r}$, which capture most of the variance. This simplifies the equation to

$$\mathbf{p} \approx (\mathbf{C} \mathbf{\Phi}_r) \mathbf{a}_r^T, \qquad \mathbf{a} \in \mathbb{R}^{1 \times r}$$
 (8)

Then the field can be estimated from sensor data s_{se} ,

$$\hat{\mathbf{u}}^T = \mathbf{\Phi}_r \hat{\mathbf{a}}_r^T = \mathbf{\Phi}_r (\mathbf{C} \mathbf{\Phi}_r)^{\dagger} \mathbf{s}_{se}$$
(9)

where $\hat{\cdot}$ represents estimated value and † indicates the Moore-Penrose pseudo-inverse.

The QR-pivoting leverages the QR factorization of the spatial mode matrix (reduced to rank r) to identify the locations that would minimize the condition ratio of $C\Phi_r$, thus improving numerical stability. This principle, based on QR discrete empirical interpolation method (Drmac and Gugercin, 2016), was used by Manohar et al. (2018). The matrix **A** is decomposed into a unitary matrix **Q**, an upper triangular matrix **R** containing non-increasing diagonal elements, and a column permutation matrix **C** such that

$$\mathbf{A}\mathbf{C}^T = \mathbf{Q}\mathbf{R} \tag{10}$$

Here, $\mathbf{A} = \mathbf{\Phi}_r$ when p = r, and $\mathbf{A} = \mathbf{\Phi}_r \mathbf{\Phi}_r^T$ when p > r. The QR factorization is performed column by column using the Householder transformation, selecting columns that maximize stability at each step.

Due to factors such as inhomogeneous illumination or particles moving into/out of the Field of View (FOV), PIV accuracy often deteriorates near the boundaries. This limitation makes applying QR-pivoting only with data near the downstream edge less practical.

In this work, we modify the QR-pivoting approach to handle multiple signal entries from each sensor. Instead of selecting individual sensor locations, we consider entire rows of PIV data as surrogates for the time series captured by probes. This requires a modification to pivoted QR decomposition referred to as "block-pivoted QR" in this paper.

The block-pivoted QR algorithm uses the same principles as QR-pivoting but evaluates blocks of data (entire rows in the flow field) instead of individual columns (single points in the flow field). This ensures that the time-delay structure is preserved while identifying the best sensor placement for flow estimation.

3. Validation

The proposed offline optimization method is assessed with both synthetic and experimental data. Two datasets from Direct Numerical Simulation (DNS) and one from experiments are utilized. The details are reported in the remainder of this section.

From this point on, the performance of the techniques used for sensor placement is tested by the velocity reconstruction error using EPOD. During testing, we temporally downsample the velocity field to simulate non-time-resolved PIV data, while use the signal located at several point of the velocity field as the sensor data, then the result is compared to the reference timeresolved field. The error ϵ is defined by

$$\epsilon = \frac{\|\mathbf{U}_{est} - \mathbf{U}_{ref}\|_2}{n_{test} \times n_c n_p} \frac{1}{U_0}$$
(11)

where the EPOD estimated velocity field \mathbf{U}_{est} and reference velocity field \mathbf{U}_{ref} from the simulation or experiment are matrices of the size $n_{test} \times n_c n_p$, n_{test} is the number of snapshots used for testing, and the definition of n_c and n_p follows Sec. 2.1. The $||\mathbf{U}||_2$ is an entry-wise Frobenius norm for an $n_{test} \times n_c n_p$ matrix \mathbf{U} , which follows

$$\|\mathbf{U}\|_{2} = \left(\sum_{i=1}^{n_{test}} \sum_{j=1}^{n_{c}n_{p}} |u_{ij}|^{2}\right)^{1/2}$$
(12)

where u_{ij} is one of the entries of U. Finally, the error is normalized by the bulk velocity or free stream velocity U_0 .

The optimal sensor positions from probes time series, row surrogate sensors, row sensors masked, block-pivoted QR in Sec. 2 as well as equidistant sensor placement will be discussed in this paper on the result from the cases below.

3.1. Fluidic pinball

The first dataset is from a 2D DNS of the wake of a *fluidic pinball* (Deng et al., 2020). The pinball setup consists of three cylinders with diameter *D*, arranged in an equilateral triangle configuration with a side length of 1.5*D*. The triangle is oriented with one vertex pointing upstream. The Reynolds number is set equal to 130, based on the freestream velocity and cylinder diameter. These conditions fall within the chaotic regime, as illustrated by Deng et al. (2020).

The flow field extends from x = 1D to x = 7D and from y = -3D to y = 3D, where x and y denote the streamwise and crosswise directions respectively. The velocity data from the original DNS mesh has been interpolated onto a Cartesian grid with a spacing of 0.08D, producing 76×76 vector fields to simulate the PIV fields.

All combinations of three probes time series, row surrogate sensors, or masked row surrogate sensors are tested. A total of 26 potential positions, uniformly spaced 0.24D apart in the *y*-direction, is explored. This results in 2600 combinations. The probe data only considers streamwise component of velocity. When simulating flow estimation with time-delay embedding, the time segment is set to 80 steps with a time spacing of 0.08D, thus resulting in a horizon slightly larger than the convection

time through the domain. The dataset is split into training (4685 non-time-resolved snapshots) and testing (4900 time-resolved frames).

3.2. Turbulent channel flow

A more challenging test case is based on the DNS of a turbulent channel flow from the Johns Hopkins Turbulence Databases. This dataset is characterized by higher spectral richness, making the EPOD estimation more difficult compared to the shedding-dominated wake flow of the previous pinball test case. The DNS for this dataset is solved in a domain of size $8\pi h \times 2h \times 3\pi h$, where *h* is the half-channel height, with a resolution of $2048 \times 512 \times 1536$ nodes, and at a Reynolds number $Re_{\tau} \approx 1000$, given the bulk velocity $U_b = 0.99994$, friction velocity $u_{\tau} = 4.9968 \times 10^{-2}$ and viscosity $v = 5 \times 10^{-5}$. The reader is referred to Li et al. (2008) for further details on the simulation.

Similarly to Discetti et al. (2018), sub-domains extending from the wall to the centreline of the channel, and with a streamwise length of h, are extracted from the full domain. The advantage of this approach is that the dataset size can be increased by exploiting statistical homogeneity in the spanwise and streamwise direction.

We are going to consider four different scenarios: (1) standard two-component (2C) planar PIV; (2) two-dimensionthree-component (2D-3C) stereoscopic PIV; (3) volumetric three-component (3C) measurements in a volume with aspect ratio, representing the ratio of the volume depth to h, of 1/8 and (4) 1/2 to simulate tomographic PIV. The 2D sub-domains include 88 × 88 velocity vectors with a grid spacing of 0.0114h, and the three-dimensional (3D) sub-domains have an aspect ratio from 0 to 1/2. The training data are independently sampled in space and time to minimize correlation between each other. Leveraging translational symmetry, this study assumes that the streamwise coordinate spans only between 0 and h, although the actual domains considered are located at different streamwise positions. The y direction ranges between 0 and h, with 0 being the position of the wall.

The probe series contains only the streamwise component of velocity. The time spacing for probe data with time series is 0.0065, and segments of 152 samples are taken at each location, thereby covering the convection throughtime of one subdomain length. The 30 potential probe positions are uniformly distributed along the downstream edge of the domain for the cases of 2D fields. For the case of 3D domain, the probes are placed along the downstream edge of the mid-plane in the spanwise direction, with the same spacing as the 2D case. Considering that 3 probes are used for the reconstruction, this results in 4060 combinations. 7200 non-time-solved snapshots are used in training and 3600 time-resolved frames are used for testing for all scenarios.

3.3. 3D experiment of propeller wake

The sensor positioning is also validated in a 3D experimental case in the wake of a windmilling propeller model. The experiment was carried out in the water tunnel at Universidad Carlos



Figure 2: Sketch of the experimental setup of tomographic PIV measurement in the wake of a windmilling propeller.

III de Madrid. A 3D printed powerless 6-blade propeller with a hub-to-tip radius r = 55 mm was installed on a pylon with null angle-of-attack with respect to the free stream, as shown in Fig. 2. The blades are driven by the water flow at a far-field velocity $U_{\infty} = 0.15$ m/s, reaching a rotation speed of about 6.0 RPM. The Reynolds number and the Strouhal number based on the diameter are 16500 and 0.22, respectively.

A time-resolved thin tomographic PIV is performed in the wake region, covering a volume of $90 \times 90 \times 7 \text{ mm}^3$, at a distance of 20 mm from the propeller axis. The flow is seeded with polyamide particles with 56 μ m diameter. Three Andor sCMOS cameras, each with 5.5 Mpixels, are deployed and operated at 80 Hz after sensor cropping. A 3 W green continuous laser pointer is used for illumination. The exposure time of the cameras is set to 2 ms. The cameras are equipped with 50 mm objectives, with 0.5× fisheye lens. The *f*[#] is set to 8 for the camera in forward scatter, and to 5.6 for the other cameras, to obtain similar imaging conditions.

The particle images are preprocessed with a slidingminimum subtraction on a 7×7 pixel kernel and Gaussian smoothing on a 3×3 pixel kernel with a standard deviation equal to 0.75. The images were then processed using a custom-made tomographic PIV software from Universita' di Napoli Federico II (Discetti et al., 2013). First, a volumetric self-calibration (Wieneke, 2008) is carried out to reduce the calibration error. The volumetric reconstruction is carried out with a multiresolution process (Discetti and Astarita, 2012a). The reconstruction algorithm is based on the camera-simultaneous multiplicative algebraic reconstruction technique (cSMART). The cSMART is a modified version of the SMART procedure proposed by Atkinson and Soria (2009) which uses the cameras sequentially. The process consists of 3 CSMART iterations on a 2x binned configuration, 3 CSMART iterations and further 3 SMART iterations on the final resolution of 11 voxels/mm. The reconstructed distributions are then interrogated with a multipass 3D cross-correlation (Elsinga et al., 2006) based on direct sparse correlations (Discetti and Astarita, 2012b). The final interrogation spot is set to 32^3 voxels, with an overlap of 75%. The vector spacing in each direction is 727 μ m.

A non-time-resolved dataset is generated by randomly selecting 4800 snapshots among the 13620 snapshots of the original sequence, and 500 frames are used for testing. The associated probe data are extracted from the full data sequence at the original sample rate from the downstream part of the PIV region on the middle plane of the volume span. The portion of tomographic PIV snapshots used for flow estimation contains $85 \times 50 \times 7$ vectors. Similarly to the synthetic-test case of the channel flow, 85×50 subdomains of 2C and 3C velocity fields are extracted for testing to simulate planar PIV and stereoscopic PIV configurations.

4. Results

4.1. Fluidic Pinball

The offline optimization process is assessed first for the case of the fluidic pinball. As discussed in Sec. 2, the map obtained by stacking the correlation of row data with their respective downstream edge point (as in Eq. 4), is shown in Fig. 3. The thick red line locates the masking threshold. The region with the correlation value under the threshold is blanked for the case of estimation with masked row sensors.

The POD spectrum is presented in Fig. 4. The spectrum reports the squared diagonal elements σ_i^2 of the diagonal matrix Σ in Eq. (1), which are related to the variance content corresponding to the *i*-th mode (Berkooz et al., 1993). The figure shows that the POD spectrum is quite compact, as expected for a low-Reynolds-number shedding-dominated flow. The first two modes contain already 60% of the energy and correspond to the vortex shedding. Ten modes are sufficient to reconstruct 90% of the energy. For this test case, it is expected that a small number of sensors should be able to achieve reasonably good reconstruction accuracy.

The positional frequency distribution on Fig. 5 represents the best 24 sensor position combinations of 3 probes, according to the reconstruction error defined by Eq. 11. To give more relevance to the most performing positioning, the combinations are sorted according to their reconstruction accuracy, and their contribution to the positional frequency distribution is weighted. The weights are set according to the following relation:

$$W = e^{-\frac{i-1}{N-1}}$$
(13)

with i representing the rank of the sensor position combination after sorting by the error value, and N the number of combinations used to generate the distribution (24 in all the cases used in this paper). Four preferred locations were identified from all types of sensors: two symmetrically located outside the wake and two on the edge of the wake. The positional distribution of row sensors is reasonably well aligned with that of the probes time series, even though the two outmost peaks are slightly shifted inwards. This suggests that the row sensors proposed for the offline optimization are reasonably good surrogates for the placement of the physical probes. Additionally, further improvement seems to be achieved by the masking, since the peaks from masked row sensors lay closer to those from point probes time series.

The figure also includes the best combination according to the block-pivoted QR positioning. The selected positions are mostly concentrated towards the inner part of the wake, thus missing mapping the outer regions. It can be expected that this greedy optimization method will be less performing when reconstructing the fields.

The 2D distribution of the estimation error, as defined in Eq. 11, is shown in Fig. 6. Compared to the 1D histogram, the 2D scatter map more effectively illustrates the correlation between different sensor formations. In these scatter maps, each spot represents a specific combination of sensor locations, with the x-axis and y-axis indicating the estimation error. The likelihood of data corresponding to specific x and y positions could be inferred from the density of points in those areas. To help readability, the scatter plot has been colour-coded for the probability density function. Here, the density ρ is counted by dividing the whole map into 100×100 regions and counting the number of spots, then normalized by the total number of spots thus $\sum_{i} \rho \Delta s_i = 1$ is complied, where the area Δs_i follows the

unit of x and y axis. For the case of row surrogate sensors (with and without masking), the error corresponds to the reconstruction error using directly the row data from each snapshot. When surrogate sensors are replaced with probes measuring time series and time-delay embedding is used in the reconstruction, the error is expected to be larger due to the lower correlation between the flow field snapshot and the probes time series.

A good qualitative indicator that offline optimization with rows is a good surrogate for sensor placement of the direct measurements is the shape of the error map. A stretched linear shape indicates that the two error metrics are interchangeable, thus the search for optimal placement can also be carried out with the row sensor. Ideally, if the leftmost point of the histogram corresponds also the downmost point, it would mean that the analysis of the reconstruction error to row sensors is able to identify also the best combination for flow estimation with probes time series.

In the case of row sensors without masking, this happens to a reasonable extent. We consider the positions obtained by selecting the three peaks of the positional frequency distribution and interpolating them with a spline to refine the positions. The positioning we obtain, corresponding to the leftmost point of the scatter plot, has an error about 19% higher than the effective optimal placement based on probes time series, as given in Table 1. The block-pivoted QR provides worse performances, as expected by Fig. 5. The concentration of probes inwards in the wake region does not provide enough details of the edge of the wake for the reconstruction.

Equidistant probe placement (thus blind to the flow distribution) works already with nearly optimal performances. This



Figure 3: Map obtained by stacking the correlation of velocity in each row with the most donwstream point of it, used for masking on the dataset of the pinball. The thick red curves bound the masked region due to a weak correlation.



Figure 4: The POD spectrum for the pinball dataset. The blue line shows the squared singular values σ_i^2 . The orange line corresponds to the cumulative sum of σ_i^2 . In both cases, the data are normalized with the sum of all the σ_i^2 .

might be ascribed to the small size of the domain in the spanwise direction, in which equidistant probe spacing allows easy capture of both the shedding structures and the wake meandering. The best performance is provided by the masked row sensors, which provide an error only 13% higher than the optimal combination of probes with time series. This is achieved using solely the snapshot PIV data, while optimization with probes time series would have required a large number of tests with probes and fields simultaneously. While this was possible in simulations, in experiments this would have been extremely onerous.

The scatter plots confirm that masked row sensors have a higher correlation to probes time series, better representing sensor locations. A sharper down-left corner in the scatter plot indicates higher correlation among preferred locations with the least velocity reconstruction error.

4.2. Turbulent channel flow

The turbulent channel flow presents the challenge of higher spectral richness. On the other hand, since Taylor's hypothesis is expected to be valid apart from the near-wall region, it might be argued that the row surrogate sensors should be better representative of the probe time series. The correlation maps for the channel case are depicted in Fig. 7. Subfigure (a) represents



Figure 5: The one-dimensional weighted positional frequency distribution of the pinball case, with the definition of W from Eq. 13. The symbols + and × represent the sensor positioning from equidistant sensors and block-pivoted QR respectively.

the correlation maps for both the 2D-2C and 2D-3C fields to simulate planar and stereoscopic PIV, as they share the same map. Subfigures (b) and (c) represent the map of thin and thick volumetric fields. In all cases, we observe a nearly monotonic decrease in correlation from downstream to upstream (in the -x-direction). The wall distance has a minor effect on the correlation in the region closer to the probe. On the other hand, towards the upstream edge, a stronger correlation is observed in the region ranging from y/h = 0 to y/h = 0.4, followed by a gradual decrease until y/h = 1. This trend is consistent across all correlation maps. The lower correlation values near the wall are likely due to the lower convective velocity and strong shear, which cause significant flow distortion before it reaches the downstream edge, as well as 3D motion. A slightly lower correlation in the channel centre might be explained by the interaction with the flow on the opposite side of the channel.

The POD spectrum for the channel flow case is illustrated in Fig. 8. Compared to the pinball case, the channel spectrum is significantly more dispersed, requiring nearly 1000 modes to capture 90% of the total energy, whereas the pinball case only needs less than 10 modes. The estimation of the temporal modes with EPOD is thus more severely challenged by the higher dimensionality of the flow. Additionally, for the cases corresponding to thick volumetric domain, the POD spectrum spreads out on an even larger number of modes.

The one-dimensional positional frequency distributions of optimal sensor positioning are shown in Fig. 9. The distribution for all aspect ratios presents 3 peaks for all types of sensors. We observe that the most performing positions for row surrogate sensors include regions located closer to the wall. This might be ascribed to the higher energy corresponding to that region. The peak positions of probes time series indicate instead that estimation with time-delay embedding is more effective for slightly larger wall-distance. This is likely due to the stronger flow distortion and 3D motion in the near-wall region, which challenge the process of time-delay embedding.

Once the correlation value is taken into account by masking low correlation region, the optimization with row surrogate

probe type	equidistant	block-	surrogate	surrogate masked	probes
field type	probes	pivoted QR	row sensors	row sensors	time series
planar PIV	0.0812	0.1176	0.0843	0.0796	0.0706

Table 1: The error of velocity reconstruction using time-delay embedding and sensors located according to different methods for the fluidic pinball. The error is defined as in Eq. 11. In the last column, the optimal case of online optimization with probes time series is included for reference.



Figure 6: The 2D probability density function of the velocity reconstruction error for probes time series versus (a) row sensors and (b) masked row sensors. The colours in the scatter plots serve as auxiliary markers, indicating the local data density. The density is defined as $\sum_{i} \rho \Delta s_i = 1$, where the Δs_i is any local area on the scatter plots.

sensor provides new locations which are closer to those of the probes time series, and thus arguably with higher potential to be more effective in the process of reconstructing once the probes are placed.

For the 3D-3C test cases, the weighted positional frequency distributions become more irregular since the prediction becomes more complex. Nonetheless, the optimal placements from row sensors and masked row sensors still have similar performance to that from probes time series as listed in Table 2. It is also worthwhile stressing that the optimal positioning from the planar field also works for the volumetric field with little degradation, which significantly reduces the computational costs of the search.

The scatter plots in Fig. 10 reflect a high similarity in velocity reconstruction between using probes time series and row surrogate sensors. The process of masking further improves this correlation, with a sharper downleft corner in almost all cases. The optimal positioning given by masked row sensors is very close to that of probes time series. This indicates that row sensors after masking are good surrogates for the probes time series.

The equidistant probes provide also reasonably good positioning for all tested cases, even slightly outperforming the row surrogate sensors positioning in planar fields. The blockpivoted QR, on the other hand, provides poorer performances, possibly because it focuses on placing probes at sites with the most information as shown in Fig. 9 (a) (i.e. region very close to the wall), rather than covering a broader area.

4.3. 3D experiment of propeller wake

The proposed method is finally tested on the velocity field from time-resolved tomographic PIV measurements of the wake of a rotating propeller. To have a solid ground truth for comparison, the original sequence is subsampled in time to produce a non-time-resolved sequence. Probes with time series are then generated at any node at the downstream edge of the domain by using data from the original time-resolved sequence. For the analysis, we consider planar data obtained by extracting the central slice of the volume, and the full volumetric measurements.

The correlation maps of the central slice and volumetric field with respect to the points on the downstream edge of the domain have similar distribution, as shown in Fig. 11. In both cases, slightly larger correlation values are observed in the wake of the hub of the windmilling propeller. The POD spectra in Fig. 12 reveal a more dispersed distribution if compared to the pinball case, thus indicating a higher level of complexity. The noise in PIV measurements further broadens the POD eigenspectrum.

equidistant	block-	surrogate	surrogate masked	probes
probes	pivoted QR	row sensors	row sensors	time series
0.0504	0.0660	0.0506	0.0496	0.0490
0.0518	0.0628	0.0516	0.0512	0.0508
0.0524	0.0632	0.0519	0.0518	0.0514
		0.0522	0.0516	
0.0570	0.0669	0.0551	0.0550	0.0549
		0.0553	0.0554	
	equidistant probes 0.0504 0.0518 0.0524 0.0570	equidistant block- pivoted QR 0.0504 0.0660 0.0518 0.0628 0.0524 0.0632 0.0570 0.0669	equidistant block- surrogate probes pivoted QR row sensors 0.0504 0.0660 0.0506 0.0518 0.0628 0.0516 0.0524 0.0632 0.0519 0.0570 0.0669 0.0551 0.0553 0.0653 0.0553	equidistant probes block- pivoted QR surrogate row sensors surrogate masked row sensors 0.0504 0.0660 0.0506 0.0496 0.0518 0.0628 0.0516 0.0512 0.0524 0.0632 0.0519 0.0518 0.0570 0.0669 0.0551 0.0550 0.0553 0.0554 0.0554

Table 2: The error of velocity reconstruction using time-delay embedding and sensors located according to different methods for the turbulent channel flow. For the volumetric test case, we consider either exploring the entire 2D downstream end of the domain (in the table as "3D positioning") or solely the central slice ("2D positioning"). The error is defined as in Eq. 11. In the last column, the optimal case of online optimization with probes time series is included for reference.

No significant differences are observed between planar and volumetric measurements.

Due to the limited field of view in the experiment, the whole domain of this case is immersed in the wake of the windmilling propeller. As shown in Fig. 13, the most frequent distributions of positions are similar for probes time series and row surrogate sensors (with and without masking). The distribution from probes time series is almost symmetric with respect to the position of the axis of the blades. We observe 4 peaks in the positional frequency distribution, although the peak located at y/r = 0.3 is weaker. On the other hand, both row sensors and masked row sensors locate the optimal sensor position near the main peaks from probes time series.

The velocity prediction errors using row sensors and masked row sensors in Table 3 are close to the ones from probes time series. As a result, the row sensors, either masked or not, demonstrate to be a good surrogate to the probe time series also in this case. This demonstrates that a preliminary experiment using planar PIV is sufficient for sensor positioning in the data-driven flow estimation even if later the experiment involves thin tomographic PIV measurements.

The scatter plots in Fig. 14 exhibit similar distributions of the reconstruction error from probes time series versus row sensors and masked row sensors in all tested configurations. This confirms that the offline optimization process is representative of the (more time consuming) online optimization process with probes spanning the volume. It must also be remarked that both types of row sensors outperform either equidistant positioning and block-pivoted QR.

Conclusions

In this paper, we introduced an offline optimization of the probe positioning for flow field estimation targeted at temporal resolution enhancement of PIV. The offline position optimization requires solely one non-time-resolved PIV sequence measurement. Under the condition that the flow is dominated by advection, time-delay embedding in the probe data can be used to enrich the dataset and improve the accuracy of the reconstruction. Our offline optimization strategy consists of using directly lines of velocity vectors in the main direction of advection from a preliminary snapshot PIV experiment, without placing any real probes, to estimate the optimal sensor placement for reconstruction with time-delay embedding. The optimal combination of probe positions is identified through a brute-force combinatorial search. The best combination is then used to locate the probes and perform the final flow estimation experiment with synchronized probe and PIV measurements.

We show that row surrogate sensors are in general capable of identifying the most promising locations for estimation of time-resolved fields once real probes are placed. The process is further enhanced by excluding regions with low correlation to potential probe positions, typically found at the downstream edge of the domain in advective flows.

Remarkably, we observed that a simple planar PIV experiment is sufficient to locate the sensor position even if the final target is time-resolution enhancement of volumetric measurements in relatively thin domains.

The offline search of the best probe positions is done by brute-force search. We have explored a modified version of a greedy optimization algorithm based on QR-pivoting to simplify this step further. Nonetheless, we observed that this method is not capable of identifying combinations of probe positions with acceptable accuracy, and it is outperformed by equidistant probe placement in all cases. We envision that further research in this direction is needed. Indeed, brute-force search is made feasible here thanks to the low computational cost of EPOD. More complex estimators might require reducing the number of estimations to perform in the offline optimization step. Solutions based on Bayesian optimization, genetic algorithms and reinforcement learning, among others, might be explored in the future. In addition, the flow estimation method might be applied to conditional measurements.

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, NEXTFLOW ERC StG). Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European

probe type	equidistant	block-	surrogate	surrogate masked	probes
field type	probes	pivoted QR	row sensors	row sensors	time series
planar PIV	0.0365	0.0375	0.0360	0.0345	0.0343
stereoscopic PIV	0.0356	0.0346	0.0345	0.0348	0.0338
thin tomographic PIV:					
with 3D positioning	0.0365	0.0378	0.0353	0.0356	0.0349
with 2D positioning			0.0363	0.0357	

Table 3: The error of velocity reconstruction using time-delay embedding and sensors located according to different methods for the experimental dataset of the propeller wake. For the volumetric test case, we consider either exploring the entire 2D donwstream end of the domain (in the table as "3D positioning") or solely the central slice ("2D positioning"). The error is defined as in Eq. 11. In the last column, the optimal case of online optimization with probes time series is included for reference.





Figure 8: The POD spectrum from the channel dataset with the sub-domain from (a) 2C planar velocity field, (b) 3C planar velocity field, (c) 3D velocity field at aspect ratio 1/8, and (d) 3D velocity field at aspect ratio 1/2.

Figure 7: Map obtained by stacking the correlation of velocity in each row with the most donwstream point of it, used to mask the two sensors in the turbulent channel flow, with different aspect ratio: (a) 2D plane, (b) 1/8 and (c) 1/2. The thick red curves bound the masked region due to a weak correlation.



Figure 9: The one-dimensional weighted positional frequency distribution of optimal sensor positioning for the channel flow across different aspect ratios of the domain, subfigures show the result from (a) 2C planar velocity field, (b) 3C planar velocity field, (c) 3D velocity field at aspect ratio 1/8, and (d) 3D velocity field at aspect ratio 1/2. The purple \times signs stand for the positioning from block-pivoted QR

Research Council. Neither the European Union nor the granting authority can be held responsible for them.

Data availability

All codes developed in this work are openly available GitHub, accessible through the link in https://github.com/erc-nextflow/sensor_placement_V1. The dataset are openly available in Zenodo, accessible at https://doi.org/10.5281/zenodo.15114116

References

- Atkinson, C., Soria, J., 2009. An efficient simultaneous reconstruction technique for tomographic particle image velocimetry. Experiments in fluids 47, 553–568.
- Bank, D., Koenigstein, N., Giryes, R., 2023. Autoencoders. Machine learning for data science handbook: data mining and knowledge discovery handbook , 353–374.
- Beresh, S.J., 2021. Time-resolved particle image velocimetry. Measurement Science and Technology 32, 102003.
- Berkooz, G., Holmes, P., Lumley, J.L., 1993. The proper orthogonal decomposition in the analysis of turbulent flows. Annual review of fluid mechanics 25, 539–575.
- Boree, J., 2003. Extended proper orthogonal decomposition: a tool to analyse correlated events in turbulent flows. Experiments in fluids 35, 188–192.
- Chaturantabut, S., Sorensen, D.C., 2010. Nonlinear model reduction via discrete empirical interpolation. SIAM Journal on Scientific Computing 32, 2737–2764.
- 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.



Figure 10: The 2D probability density function of the velocity reconstruction error using probes with time-series versus row sensors (left) or row sensors masked (right). The computation domains are, from top to bottom: 2C planar (1st row), 3C planar (2nd row), thin volumetric with aspect ratio 1/8 (3rd row), volumetric with aspect ratio 1/2 (4th row). The colour of the spots represents local density ranging from low (blue) to high (red). It is set $\sum \rho \Delta s_i = 1$, where

the Δs_i is any local area on the scatter plots.



Figure 11: The front and left view of the blades model and the correlation map for (a) planar PIV obtained by extracting data from the centre slice of the tomographic PIV domain and (b) by averaging along the depth direction the correlation map from the full tomographic PIV field. The thick curves shows the left margin of the region after masking by correlation.

- Clark, H., Naghib-Lahouti, A., Lavoie, P., 2014. General perspectives on model construction and evaluation for stochastic estimation, with application to a blunt trailing edge wake. Experiments in fluids 55, 1–19.
- 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.
- Deng, Z., Chen, Y., Liu, Y., Kim, K.C., 2019. Time-resolved turbulent velocity field reconstruction using a long short-term memory (LSTM)-based artificial intelligence framework. Physics of Fluids 31.
- Discetti, S., Astarita, T., 2012a. A fast multi-resolution approach to tomographic PIV. Experiments in fluids 52, 765–777.
- Discetti, S., Astarita, T., 2012b. Fast 3D PIV with direct sparse crosscorrelations. Experiments in fluids 53, 1437–1451.
- Discetti, S., Bellani, G., Örlü, R., Serpieri, J., Vila, C.S., Raiola, M., Zheng, X., Mascotelli, L., Talamelli, A., Ianiro, A., 2019. Characterization of verylarge-scale motions in high-Re pipe flows. Experimental Thermal and Fluid Science 104, 1–8.
- Discetti, S., Natale, A., Astarita, T., 2013. Spatial filtering improved tomographic PIV. Experiments in fluids 54, 1–13.
- 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.
- Drmac, Z., Gugercin, S., 2016. A new selection operator for the discrete empirical interpolation method—improved a priori error bound and extensions. SIAM Journal on Scientific Computing 38, A631–A648.
- Elsinga, G.E., Scarano, F., Wieneke, B., van Oudheusden, B.W., 2006. Tomographic particle image velocimetry. Experiments in fluids 41, 933–947.
- Erichson, N.B., Mathelin, L., Yao, Z., Brunton, S.L., Mahoney, M.W., Kutz, J.N., 2020. Shallow neural networks for fluid flow reconstruction with limited sensors. Proceedings of the Royal Society A 476, 20200097.
- Ewing, D., Citriniti, J.H., 1999. Examination of a LSE/POD complementary technique using single and multi-time information in the axisymmetric shear layer, in: IUTAM Symposium on Simulation and Identification of Organized Structures in Flows: Proceedings of the IUTAM Symposium held in Lyngby, Denmark, 25–29 May 1997, Springer. pp. 375–384.
- Ghil, M., Allen, M.R., Dettinger, M.D., Ide, K., Kondrashov, D., Mann, M.E., Robertson, A.W., Saunders, A., Tian, Y., Varadi, F., et al., 2002. Advanced spectral methods for climatic time series. Reviews of geophysics 40, 3–1.
- 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, 1–16.

Kramer, M.A., 1991. Nonlinear principal component analysis using autoasso-

ciative neural networks. AIChE journal 37, 233–243.

- Lee, J.A., Verleysen, M., et al., 2007. Nonlinear dimensionality reduction. volume 1. Springer.
- Li, Y., Perlman, E., Wan, M., Yang, Y., Meneveau, C., Burns, R., Chen, S., Szalay, A., Eyink, G., 2008. A public turbulence database cluster and applications to study Lagrangian evolution of velocity increments in turbulence. Journal of Turbulence 9.
- Manohar, K., Brunton, B.W., Kutz, J.N., Brunton, S.L., 2018. Data-driven sparse sensor placement for reconstruction: Demonstrating the benefits of exploiting known patterns. IEEE Control Systems Magazine 38, 63–86.
- Meilă, M., Zhang, H., 2024. Manifold learning: What, how, and why. Annual Review of Statistics and Its Application 11.
- Moreno Soto, Á., Güemes, A., Discetti, S., 2024. Complete flow characterization from snapshot PIV, fast probes and physics-informed neural networks. Computer Methods in Applied Mechanics and Engineering 419, 116652.
- Nakai, K., Yamada, K., Nagata, T., Saito, Y., Nonomura, T., 2021. Effect of objective function on data-driven greedy sparse sensor optimization. IEEE Access 9, 46731–46743.
- Nickels, A., Ukeiley, L., Reger, R., Cattafesta III, L., 2020. Low-order estimation of the velocity, hydrodynamic pressure, and acoustic radiation for a three-dimensional turbulent wall jet. Experimental Thermal and Fluid Science 116, 110101.
- van Oudheusden, B., 2013. PIV-based pressure measurement. Measurement Science and Technology 24, 032001.
- Papadopoulou, M., Raphael, B., Smith, I.F., Sekhar, C., 2014. Hierarchical sensor placement using joint entropy and the effect of modeling error. Entropy 16, 5078–5101.
- Paris, R., Beneddine, S., Dandois, J., 2021. Robust flow control and optimal sensor placement using deep reinforcement learning. Journal of Fluid Mechanics 913, A25.
- Sicot, C., Perrin, R., Tran, T., Borée, J., 2012. Wall pressure and conditional flow structures downstream of a reattaching flow region. International journal of heat and fluid flow 35, 119–129.
- Sirovich, L., 1987. Turbulence and the dynamics of coherent structures. I. Coherent structures. Quarterly of applied mathematics 45, 561–571.
- Tinney, C., Coiffet, F., Delville, J., Hall, A., Jordan, P., Glauser, M., 2006. On spectral linear stochastic estimation. Experiments in fluids 41, 763–775.
- 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,





Figure 12: The POD spectrum from the blade dataset with the sub-domain from (a) planar velocity field, (b) 3C planar velocity field, (c) thin tomographic field.

Figure 13: The one-dimensional weighted positional distribution of optimal sensor positioning after testing with row sensors, row sensors masked and probes with time-series: (a) 2C planar PIV; (b) 3C planar PIV; (c) tomographic PIV field. The purple \times signs stand for the positioning from block-pivoted QR.



Figure 14: The scatter plots of velocity reconstruction error using probes with time-series versus using row sensors (left) and row sensors masked (right). The computation domains are, from top to bottom: 2D-2C field in the central plane (top); 2D-3C field in the central plane (central); thin volumetric field (bottom). The colour of the spots represents local density ranging from low (blue) to high (red). The +, × and ★ show the result from the placement of equidistant, block-pivoted and optimal combination tested by row sensors or row sensors masked. The density is defined as $\sum_{i} \rho \Delta s_i = 1$, where the Δs_i is any local area on the scatter plots.

1-20.

