



# Outlier detection for PIV statistics based on turbulence transport

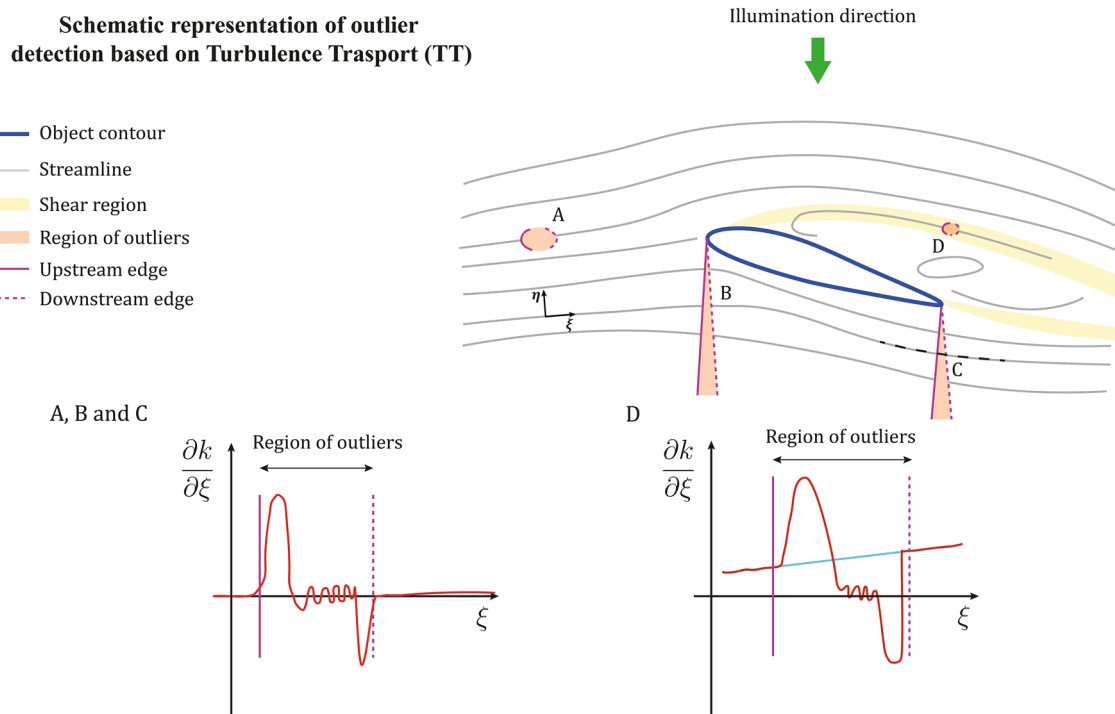
E. Saredi<sup>1</sup> · A. Sciacchitano<sup>1</sup> · F. Scarano<sup>1</sup>

Received: 31 August 2021 / Revised: 20 October 2021 / Accepted: 9 December 2021 / Published online: 23 December 2021  
© The Author(s) 2021

## Abstract

The occurrence of data outliers in PIV measurements remains nowadays a problematic issue; their effective detection is relevant to the reliability of PIV experiments. This study proposes a novel approach to outliers detection from time-averaged three-dimensional PIV data. The principle is based on the agreement of the measured data to the turbulent kinetic energy (TKE) transport equation. The ratio between the local advection and production terms of the TKE along the streamline determines the admissibility of the inquired datapoint. Planar and 3D PIV experimental datasets are used to demonstrate that in the presence of outliers, the turbulent transport (TT) criterion yields a large separation between correct and erroneous vectors. The comparison between the TT criterion and the state-of-the-art universal outlier detection from Westerweel and Scarano (Exp Fluids 39:1096–1100, 2005) shows that the proposed criterion yields a larger percentage of detected outliers along with a lower fraction of false positives for a wider range of possible values chosen for the threshold.

## Graphical abstract



✉ E. Saredi  
e.saredi@tudelft.nl

<sup>1</sup> Faculty of Aerospace Engineering, TU Delft, Delft, The Netherlands

## 1 Introduction

In particle image velocimetry (PIV), outliers are spurious vectors that exhibit large *unphysical variations* in magnitude and direction from neighbouring valid vectors (Westerweel 1994). Despite the above definition, most proposed and used approaches for outliers detection are not based on flow physics, but rather on statistical data analysis. The latter stems from the open question as to what aspect of the flow physics should be considered to judge a velocity vector as unphysical. Furthermore, planar data are often unsuited to the use of physical principles, given that the constitutive flow equations consider the three physical dimensions of space. Among the factors leading to the occurrence of an outlier are light reflections from solid surfaces or background, poor illumination, shadow regions, inhomogeneous flow seeding, inadequate image recording and interrogation (Hart 2000; Lazar et al. 2010; Sciacchitano 2019; among others). Most methods presented in the literature focus on the detection of outliers in the instantaneous flow field (Higham et al. 2016), and the state-of-the-art approach is based on the universal outlier detection (UOD, Westerweel and Scarano 2005), which generalises the median detector presented earlier by Westerweel (1994). Duncan et al. (2010) further extended the UOD approach for unstructured data such as those obtained by particle tracking velocimetry (PTV) or spatially adaptive interrogation algorithms. Masullo and Theunissen (2016) dealt with the relevant problem of clusters of false vectors and proposed to combine a spatial coherence estimator with a Gaussian-weighted distance-based averaging median.

Proper orthogonal decomposition (POD) was also considered as a basis to detect outliers (Wang et al. 2015; Raiola et al. 2015; Higham et al. 2016), using a low-order reconstruction to detect outliers. The latter approach was proven effective also in detecting clustered outliers, compared to the UOD.

An early attempt to use constitutive equations for outlier detection is due to Song et al. (1999) who invoked compliance to the continuity equation, numerically evaluated from a Delaunay tessellation of the domain. For experiments in the incompressible flow regime, outliers would produce unphysical nonzero values of the velocity divergence. However, a unique value of the threshold that separates outliers from other acceptable sources of error (e.g. small-amplitude measurement noise) could not be identified, as deductible from the data presented by Azijli and Dwight (2015).

The aforementioned outlier detection methods are typically effective for the instantaneous velocity fields where a single vector or a cluster thereof largely departs in

magnitude and direction from the neighbouring points. Such outliers need to be detected and replaced, or omitted, when estimating statistical flow properties like the mean value and its fluctuations. If not, they yield erroneous flow statistics, whereby the outliers depart less markedly from the neighbouring vectors and therefore are not easily detectable with state-of-the-art outlier detection approaches.

The advent of three-dimensional PIV techniques (Rafael et al. 2018) and in particular of particle-based analysis (Shake-The-Box, STB, Schanz et al. 2016) has established new approaches to evaluating the flow statistics from experiments. The instantaneous flow velocity is obtained directly at the sparse positions occupied by the particle tracers. Therefore, ensemble statistics are obtained partitioning the domain into cubic voxels (or bins) arranged on a Cartesian grid and evaluating locally the ensemble average and fluctuations from all particles velocity belonging to the considered voxel. Techniques that perform accurately this operation have been devised by Agüera et al. (2016) and more recently by Godbersen and Schröder (2020).

The occurrence of outliers in PTV data is notably more frequent than in data produced with cross-correlation analysis (2nd PIV challenge, Stanislas et al. 2005), thus bringing forward the problem of outliers detection from ensemble statistics of 3D-PTV data.

In the present work, the principle of using constitutive equations for the detection of outliers from statistical datasets is revised, invoking the turbulence transport equation as a consistent choice. This article first recalls the terms of turbulence transport (Sect. 2.1) and then illustrates how it can be invoked to isolate data that violate the principle (Sect. 2.2). The validity of the approach is examined with three experimental datasets in Sect. 3. The effectiveness and reliability are compared with the state-of-the-art UOD in Sect. 3.3.

## 2 Turbulent transport detection criterion

### 2.1 Working principle

Let us consider a statistical dataset whereby the measured velocity is represented by means of the Reynolds decomposition into the mean and fluctuating part. Taking for instance the streamwise velocity component  $u$ , such decomposition reads as:

$$u = \bar{u} + u' \quad (1)$$

The turbulent kinetic energy (TKE, here denoted by the symbol  $k$ ) considers the statistical fluctuations from all velocity components:

$$k = 1/2(\overline{u'^2} + \overline{v'^2} + \overline{w'^2}) \tag{2}$$

and its behaviour is governed by the turbulence transport equation (Hinze 1975):

$$\frac{Dk}{Dt} = T + P - \epsilon \tag{3}$$

The above expression states that the variation of TKE of a fluid parcel is governed by transport  $T$ , production  $P$  and dissipation  $\epsilon$ . The main question raised here is to what extent can the occurrence of a statistical outlier be detected invoking the above physical principle. For further use, the terms of Eq. 3 are explained below:

$$\frac{Dk}{Dt} = \frac{\partial k}{\partial t} + A \text{ material derivative of } k, \tag{4}$$

$$A = \bar{u}_i \frac{\partial k}{\partial x_i} \text{ advection,} \tag{5}$$

$$P = -\overline{u'_i u'_j \frac{\partial \bar{u}_i}{\partial x_j}} \text{ production,} \tag{6}$$

$$T = -\frac{1}{2} \frac{\partial \overline{u'_i u'_j u'_i}}{\partial x_j} + \nu \frac{\partial^2 k}{\partial x_j \partial x_j} - \frac{1}{\rho} \frac{\partial \overline{u'_j p'}}{\partial x_i} \tag{7}$$

transport by turbulence and pressure,

$$\epsilon = \nu \overline{\frac{\partial u'_i}{\partial x_j} \frac{\partial u'_i}{\partial x_j}} \text{ dissipation.} \tag{8}$$

From a recent review of studies describing the TKE budget (Ikhennicheu et al. 2020), in turbulent shear flows, a first approximation of the behaviour for the turbulent quantities is that of quasi-equilibrium turbulence, whereby the turbulent properties vary very gradually along a streamline. Under this condition, the production and the dissipation are the dominant terms at the right-hand side of Eq. 2, with the former acting as the upper bound of the convection of TKE along a streamline, since a non-null dissipation reduces the TKE.

It is hypothesized here (and illustrated in the considered experiments of Sect. 3) that the presence of data outliers generally leads to an overestimation of TKE and locally biases the RMS fluctuations. This condition can be detected as unphysical, invoking a criterion based on Eq. 3 or a derived form of it as shown in the remainder.

### 2.2 Detection criterion

The interaction of outliers and turbulence transport is first illustrated schematically. Let us consider the statistical dataset from the measurements around an airfoil at incidence

(Fig. 1), which corresponds to one of the experiments presented in the remainder. The measurement yields the statistical description of the velocity field  $\bar{\mathbf{V}}$ ,  $\mathbf{V}'$  and an estimate of the velocity fluctuations  $k$ . Flow separation from the leading edge results in a large flow recirculation bounded by turbulent shear layers. The laser light refraction at the edges of the transparent airfoil produces regions with no illumination (indicated as A and B in Fig. 1). Additionally, two small regions of outliers are hypothesised (region A ahead of the airfoil and region D in the shear layer) as resulting from background light reflections.

In the above regions, a significantly more frequent occurrence of outliers is hypothesized, in turn causing a localised overestimation of the TKE.

However, let us recall that along any time-average streamline, the evolution of  $k$  must comply with Eq. 3.

Therefore, following a streamline that crosses a region of outliers, it is expected that the absolute variation of TKE along the streamline  $\left| \frac{\partial k}{\partial \xi} \right|$  will largely exceed the value allowed by Eq. 3. When outliers appear featuring a cluster, the principle is violated at its upstream and downstream edges. In the irrotational regions (A, B and C of Fig. 2)  $k$  is expected to remain constant and close to null (laminar shear free flow), alongside its spatial derivative along the streamline, i.e.  $\left| \frac{\partial k}{\partial \xi} \right| = 0$  (light blue line in Fig. 2). The occurrence of outliers in the measurement will produce a visible increase of  $\left| \frac{\partial k}{\partial \xi} \right|$  at the upstream and downstream edges. Taking into consideration some measurement noise for a threshold level (black dashed line in Fig. 2), a single value of  $\left| \frac{\partial k}{\partial \xi} \right|$  may allow detecting the data at the upstream boundary as erroneous.

In the regime of turbulent shear (region D in Fig. 2-right) instead, the variation of  $k$  along the streamline is not expected to be null, with the term  $P$  dominating the right-hand side of Eq. 3. A threshold must be based on an estimate of the local production of turbulence, which requires the measurement of the shear rate and the Reynolds stress tensor.

Based on the above discussion, one can define a criterion that states the admissible value for TKE advection along a streamline, compared to an estimation of TKE production:

$$\rho_{TT} = \frac{|A|}{|P|} \tag{9}$$

Such admissible values of  $\rho_{TT}$  are to be of  $O(1)$  or smaller. Moreover, considering that the estimation of  $P$ , performed considering the unscrutinised velocity field, is affected by measurement noise, the production term in Eq. 9 is replaced with the median  $P_m$  of the adjacent vectors in a kernel  $5 \times 5$ , in analogy to Westerweel and Scarano (2005). Similarly, to avoid the ratio losing

significance when the production is null (for instance in regions of laminar flow, free of shear), an additive term  $\gamma_{TT}$  is included on the denominator that scales with the uncertainty on the numerator  $A$  of Eq. 9.

With the latter modifications, Eq. 9 becomes:

$$\rho_{TT} = \frac{|A|}{|P_m| + \gamma_{TT}} < 5 \tag{10}$$

Due to the statistical nature of the terms of Eq. 10, it must be underlined that the presented approach is only meant for flow statistics and cannot be directly applied to instantaneous velocity fields.

A suitable value for the term  $\gamma_{TT}$  can be estimated assuming that the velocity is measured with uncertainty not exceeding 5% of a reference velocity, e.g.  $V_\infty$ :

$$\gamma_{TT} = \frac{(0.05V_\infty)^3}{\Delta} \tag{11}$$

with  $\Delta$  representing the pitch of the grid where the velocity vector field is evaluated.

Equation 10 is written already in the form of a detection criterion, whereby a threshold value of 5 is proposed here that separates admissible (unit order of magnitude) from non-admissible (higher order of magnitude) values of  $\rho_{TT}$ . The above choice is scrutinised with the application of the criterion to measured data from wind tunnel experiments.

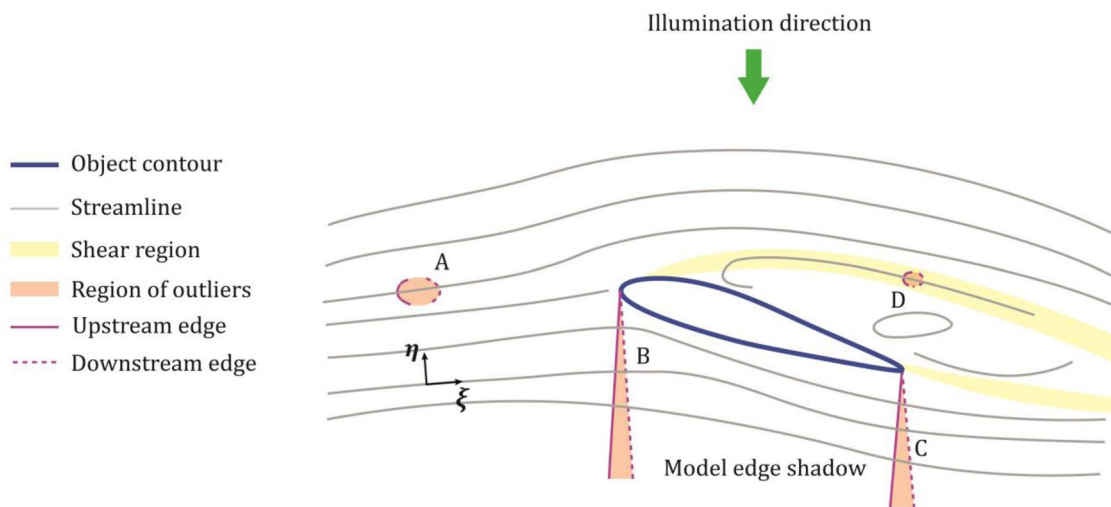
### 3 Experimental assessment

#### 3.1 Selected datasets

Three experimental datasets have been considered: the 2D velocity field around a NACA 0012 airfoil, the near-wake of a truncated cylinder and that of the Ahmed body (Ahmed et al. 1984).

Planar, two-component PIV measurements are taken for the airfoil case. The flow is seeded with water–glycol droplets of 1  $\mu\text{m}$  median diameter. Illumination is provided by a Quantel *Evergreen 200* Nd:YAG laser ( $2 \times 200$  mJ at 15 Hz), and the illuminated particles are imaged by a LaVision *Imager sCMOS* camera ( $2560 \times 2160$  pixels, 16 bits). A description of the setup can be found in Adatrao et al. (2021). Figure 3a shows an instantaneous recording featuring, aside from uniform seeding concentration, shadow region emanating from the leading edge. Multi-pass, 2D cross-correlation analysis is used to evaluate the particles motion.

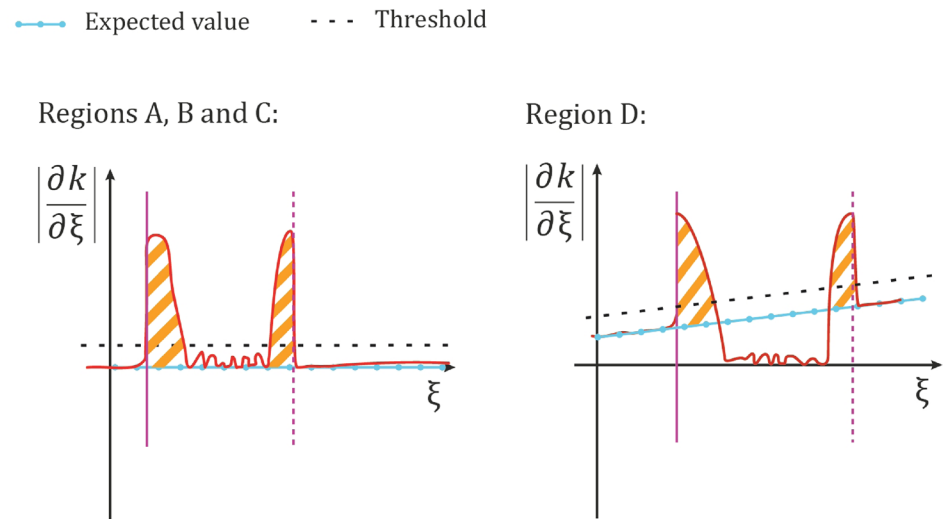
For the near-wake of the cylinder, a tomographic PIV system was used. In order to achieve a measurement volume of 6 L, sub-millimetre helium-filled-soap-bubbles (HFSB) have been used as tracer particles (Faleiros et al. 2019; Scarano et al. 2015). The illumination was provided by a Quantronix *Darwin Duo* Nd:YLF laser ( $2 \times 25$  mJ at 1 kHz). The imaging system comprised four high-speed Photron *FAST-CAM SAI* cameras ( $1024 \times 1024$  pixels, 10 bits) featuring a tomographic system with a large aperture (Schneiders et al. 2016). The region of interest is illuminated from downstream, and the round shape of the cylinder surface produces



**Fig. 1** Flow around an airfoil at incidence. The separated region at the suction side is bounded by free shear layers. Outlier regions are highlighted in light red: A, B and C in the external flow around the

airfoil; D inside the shear layer. Red continuous and dashed lines mark the upstream and downstream edges of each region

**Fig. 2** Schematic representation of the expected variation of turbulent kinetic energy along a streamline. At the boundaries of the erroneous region, an unphysical increase of TKE flags erroneous vectors



a strong reflection towards two of the cameras, placed at the same height of the model. A raw image with such reflection is shown in Fig. 3b. The recordings were analysed with the STB algorithm (Schanz et al. 2016), yielding the particles velocity along their trajectory. Ensemble-average over cubic bins was performed returning the time-average velocity and its fluctuations on a Cartesian grid.

The near-wake of the Ahmed body was investigated by Saredi et al. (2020) using robotic volumetric PIV (Jux et al. 2018). HFSB were used to seed the flow stream. The technique makes use of a coaxial volumetric velocimeter (Schneiders et al. 2018), whereby the laser light propagates approximately along the same direction as that of the four CMOS imagers composing the LaVision *Minishaker* (1920 × 1200 pixels, 10 bits). As a result, the recorded images tend to collect light reflected from objects and the illuminated background. The situation is illustrated in Fig. 3c. The particle motion is obtained with STB in a way similar to the previous case.

The main experimental parameters of the three considered datasets are listed in Table 1.

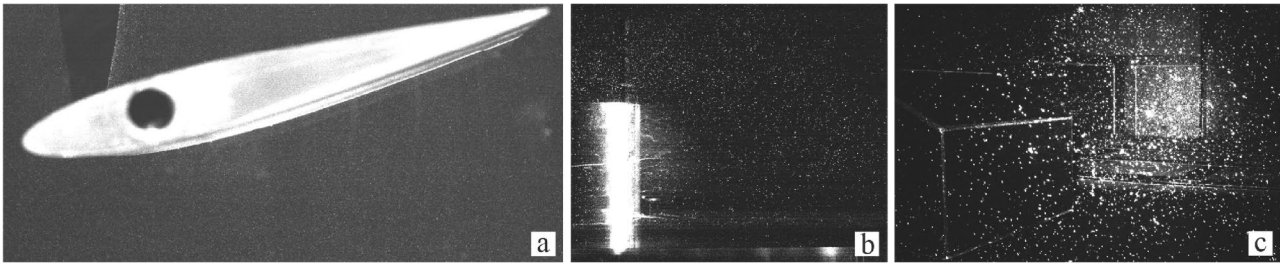
### 3.2 Velocity field statistics

The error sources described above introduce outliers at random time instants. This process is illustrated in the example of Fig. 4 in the left column, where the instantaneous measurement features isolated outliers and, more importantly, regions of finite extent with a large fraction of unphysical data. The cross-correlation analysis fails in the shadow region above the airfoil, which leads to random velocity vectors. For the cylinder case, the strong reflection on the surface of the body causes the appearance of erroneous particle tracks along one of the tomographic lines of sight.

These tracks feature high-amplitude errors. The unphysical values extend also inside the solid object. The background reflection in the case of the Ahmed body wake introduces erroneous tracks in the free-stream region. Also in this case, this region extends along the viewing direction. The outliers that appear in the instantaneous fields are markedly departing from the correct measurements. When the statistical dataset is built to evaluate the time-average velocity field (Fig. 4, centre-column), such departure becomes less evident as a result of the statistical averaging between spurious and correct data. For the 2D-wing case, the region corresponding to both the shadow and the perspective section of the wing is characterized by erroneous low velocity. For the cylinder case, laser reflections generate invalid vectors inside the cylinder geometry. Furthermore, a region of flow characterized by quasi-null streamwise velocity can be spotted in the area corresponding to the erroneous tracks. Such areas are highlighted with dashed ellipses in Fig. 4. Similarly, the presence of spurious tracks in the near-wake of the Ahmed body produces an unphysical region of deceleration and acceleration at  $X=0.2$  m in the free-stream region of the ensemble-averaged flow field. The effect of instantaneous velocity outliers affects even more markedly the RMS velocity fluctuations (Fig. 4, right column). This effect is illustrated here by displaying the quantity  $\sqrt{k}/V_\infty$ . All outlier regions feature a sharp increase of TKE with respect to the surrounding. It can be noted that the extent and the shape of the outlier regions is more visible for the velocity fluctuations rather than for the time-averaged velocity.

However, the TKE alone can hardly be used as a criterion for outliers detection, because its values are not bounded by criterion and the threshold for admissible values varies greatly among experiments. For instance, while for the





**Fig. 3** Raw images from the three experiments. **a** NACA 0012 airfoil; **b** near-wake of a truncated cylinder; **c** near-wake of the Ahmed body

airfoil case,  $\frac{\sqrt{k}}{V_\infty} < 0.4$  would be an appropriate criterion to identify the entire erroneous data caused by the shadow, the same threshold would introduce false positives (i.e. flag correct measurements as erroneous) in the shear region of the cylinder wake.

### 3.3 Comparison with UOD

The effectiveness of the turbulence transport-based criterion is compared to the universal outlier detection (UOD, West-erweel and Scarano 2005). In the latter work, outliers are identified by a median-based residual  $r_0^*$  defined as:

$$r_0^* = \frac{|V_0 - V_m|}{r_m + \gamma_{UOD}} \quad (12)$$

where  $V_0$  is the considered velocity vector,  $V_m$  the median value across its neighbourhood,  $r_m$  the median of the neighbours residuals defined as  $r_i = |V_i - V_m|$ , and  $\gamma_{UOD}$  is the minimum normalization level, often set to 0.1 pixels. The UOD criterion has been postulated, based on experiments as  $r_0^* > 2$  and the same threshold value is used in the present analysis.

The application of the two outliers indicators to the present data, namely  $\rho_{TT}$  of Eq. 10 and the residual  $r_0^*$  of Eq. 12, is presented in Fig. 5. The condition  $r_0^* > 2$  does not detect false vectors for the airfoil and the cylinder case. In

the former, only a small region at the leading edge of the shadow would be considered erroneous. In the latter, the area affected by reflection is not detected as erroneous, with only some vectors at the edges of the flow domain indicated as outliers. The same criterion does detect erroneous vectors in the Ahmed body wake. However, also correct measurements in the shear layers are labelled as outliers (false positives).

In the right column of Fig. 5 the  $\rho_{TT}$  criterion is displayed. The highest value is attained at the edges of the erroneous regions, ( $\rho_{TT} > 10$ ), well separated, by approximately one order of magnitude, with respect to the adjacent correct measurements. The edges of the airfoil shadow region and the edges of the area blocked by perspective view are clearly detected as false vectors. The edges of the erroneous region in the cylinder wake are also clearly detected, again with a large separation to the correct portion of the flow where  $\rho_{TT} \sim 1$ .

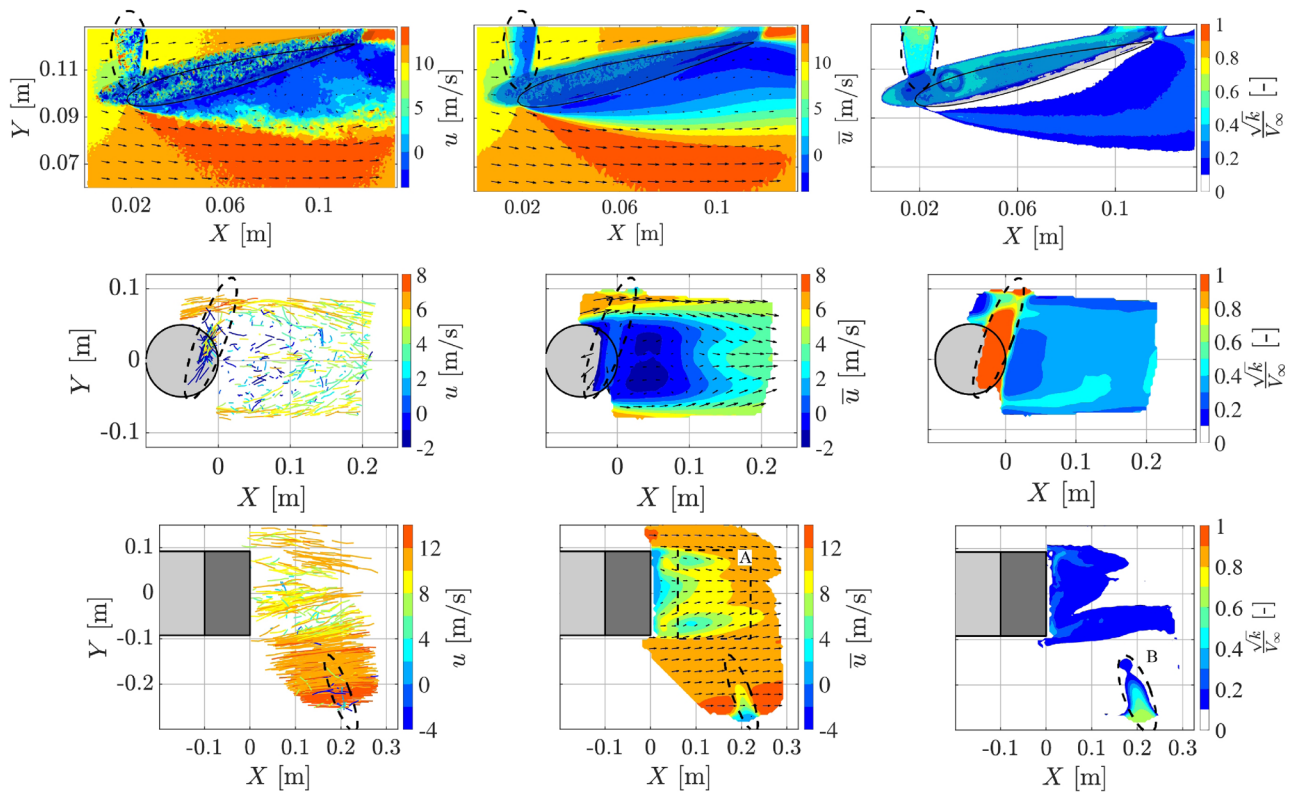
In the Ahmed body wake,  $\rho_{TT}$  correctly detects the boundary of the erroneous region highlighted by the dashed ellipse in Fig. 5 (right). Two regions at the back edge of the Ahmed body are also highlighted as erroneous, suggesting the presence of false positives for  $\rho_{TT}$  too, in those regions.

### 3.4 Detection ratio and false positive

The performance of the above detection methods is assessed introducing the detection and false positive rates  $\eta_d$  and  $\eta_{fp}$ ,

**Table 1** Measurement parameters for the three considered experiments

Experiment	NACA 0012	Near-wake truncated cylinder	Near-wake Ahmed body
2D/3D	2D	3D	3D
Number of snapshots	1000	2000	10,000
Seeding	Fog droplets ( $d_p = 1 \mu\text{m}$ )	HFSB ( $d_p = 300 \mu\text{m}$ )	HFSB ( $d_p = 300 \mu\text{m}$ )
Illumination	Nd:YAG laser ( $2 \times 200 \text{ mJ @ } 15 \text{ Hz}$ )	Nd:YLF laser ( $2 \times 25 \text{ mJ @ } 1 \text{ kHz}$ )	Nd:YLF laser ( $2 \times 25 \text{ mJ @ } 1 \text{ kHz}$ )
Recording device	sCMOS ( $2560 \times 2160 \text{ px}^2$ , 16 bits, $6.5 \mu\text{m}$ pixel pitch)	CMOS ( $1024 \times 1024 \text{ px}^2$ , 12 bits, $20 \mu\text{m}$ px pitch)	CMOS ( $1024 \times 1024 \text{ px}^2$ , 12 bits, $20 \mu\text{m}$ px pitch)
Recording method	Double frame/single exposure	Time-resolved/single exposure	Time-resolved/single exposure
Processing technique	Cross-correlation	STB	STB



**Fig. 4** Left: Instantaneous velocity vectors or tracks. Centre: time-averaged velocity vectors; region A is used for determining  $\eta_{fp}$  (Fig. 6). Right: normalized fluctuations,  $\sqrt{k}/V_\infty$ . Outliers regions are highlighted by a dashed ellipse; region B is used to determinate  $\eta_d$

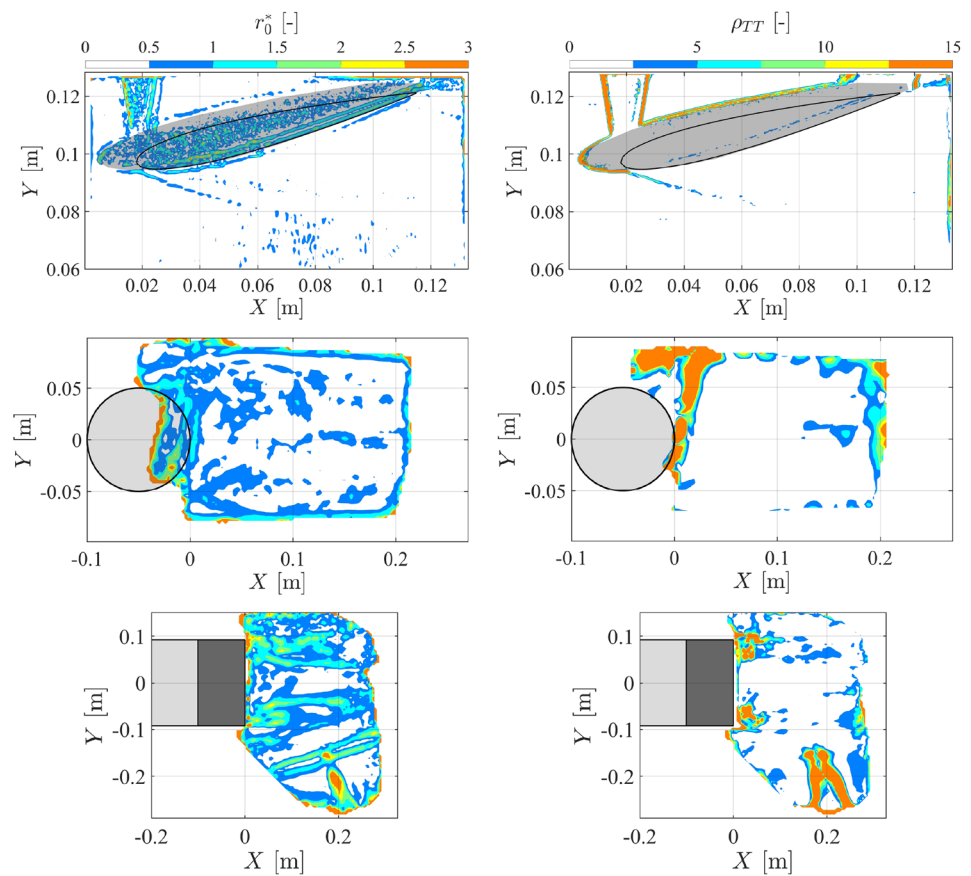
respectively. The former is defined as the number of spurious vectors correctly flagged as erroneous divided by the total number of outliers. The false positive rate  $\eta_{fp}$  is defined as the ratio of correct vectors flagged detected as outliers and the total number of correct vectors. The criterion robustness is defined by how constant the above two parameters remain by varying the threshold. It is chosen to vary the UOD criterion between 0 and 4 and  $\rho_{TT}$  between 0 and 10. A portion of the dataset in the near-wake of the Ahmed body has been considered where the presence of erroneous vectors can be clearly excluded, yet the flow exhibits significant shear and turbulence. The planar section of the considered volume is indicated as region A in Fig. 4 (centre). Instead, for the estimation of  $\eta_d$  erroneous vectors need to be known with a high level of confidence. In this case, the area circled and indicate as region B in Fig. 4 (right) has been considered. The proposed principle aims to detect erroneous vectors at the edge of correct measurement regions; only the vectors located at the edge of the faulty area have been considered. Once isolated, this group of vectors, the amount of them flagged as erroneous by the UOD and by the turbulence transport-based outlier detection has been calculated. The results obtained in terms of  $\eta_d$  and  $\eta_{fp}$  by the two methods, varying

(Fig. 6). For the 3D datasets, the data are presented in a slice crossing the outlier region ( $Z = 0.13$  m for the Ahmed body,  $Z = 0.05$  m for the cylinder)

the corresponding thresholds, are presented in Fig. 6. It has to be noted that the optimal values of  $\eta_{fp}$  and  $\eta_d$  are 0 and 1, respectively. It is expected that, when increasing the value of the threshold for both criteria, the correct detection probability decreases, alongside the false positive rate. Both criteria yield the same expected behaviour. The UOD criterion features a rapidly declining detection rate, with a value of approximately 50% for  $r_0^* = 2$ . At this point the fraction of less than 5% for the false positives is acceptably small. The turbulence transport-based criterion detects approximately 90% of the erroneous vectors with a threshold of 5, with a fraction of false positives similar to that of UOD. Moreover, the difference subtended between  $\eta_d$  and  $\eta_{fp}$  remains rather in a wide range of values for the threshold:  $\Delta\eta > 0.8$  when  $\rho_{TT} \in [1, 10]$ . In contrast, such difference is limited to 0.4 for UOD in the range  $r_0^* \in [1, 3]$ .

An additional indication of the outlier detection stability can be obtained by examining the fraction of detected outliers with respect to a variation of the threshold. The histograms of  $r_0^*$  and  $\rho_{TT}$  are shown in Fig. 7. Similarly to the results in Fig. 6, both methods yield a decreasing number of detected vectors with an increasing value of the threshold. A significant difference is observed in the slope

**Fig. 5** Left: normalized residual  $r_0^*$  of the UOD criterion (West-erweel and Scarano 2005). Right: turbulence transport ratio  $\rho_{TT}$  defined by Eq. 10. For the 3D datasets, the data are presented in a slice crossing the outlier region ( $Z=0.13$  m for the Ahmed body,  $Z=0.05$  m for the cylinder)



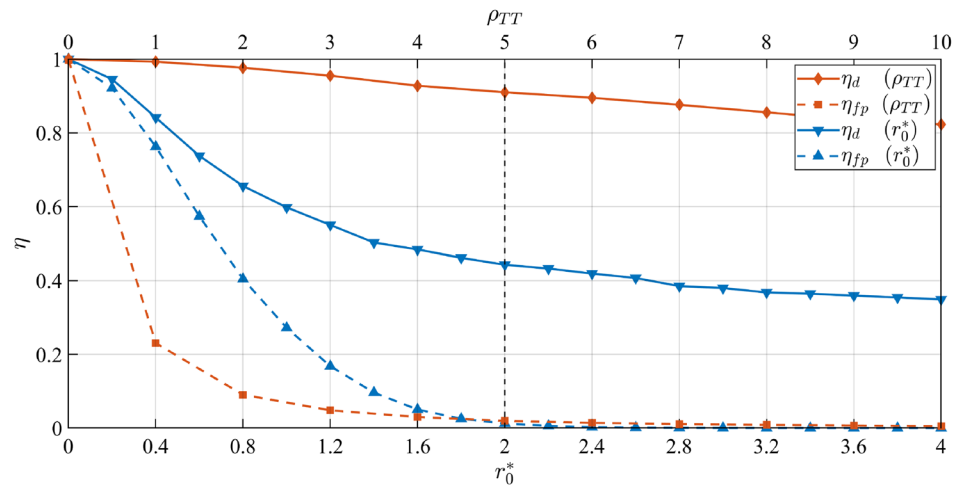
of the curves corresponding to the two criteria. While for the UOD criterion unit change of  $r_0^*$  causes a variation by a decade in the histogram (slope  $10^{-1}$ ), the same change of  $\rho_{TT}$  introduces a negligible variation of detected outliers (slope  $10^{-1/6}$ ). It is thus concluded that the detection of outliers based on the turbulence transport principle is a significantly more stable estimator of statistical outliers alongside yielding minimal occurrence of false positives.

## 4 Conclusions

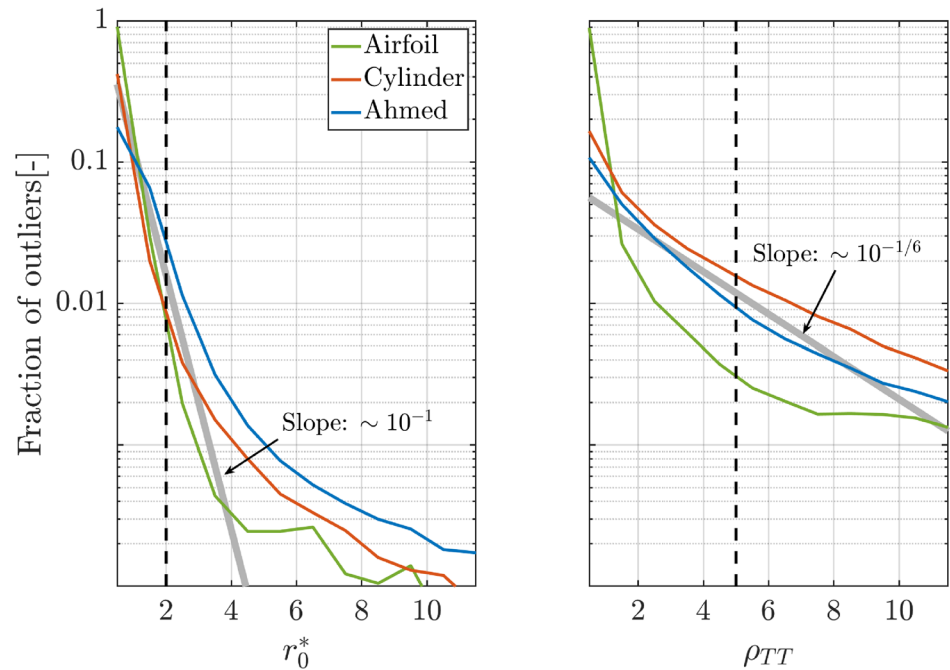
A novel outlier detection criterion is presented that invokes the physical principle of turbulence transport and applies to the flow statistics evaluated by PIV or PTV, preferably in 3D datasets. The approach relies on the fact that the measured velocity fluctuations are bound to comply with the governing equation of turbulent kinetic energy transport. The ratio



**Fig. 6** Comparison of the detection ratio  $\eta_d$  and the false positive ratio  $\eta_{fp}$  obtained by UOD and  $\rho_{TT}$ . Data extracted from Ahmed body near wake



**Fig. 7** Histogram of detected outliers by UOD (left) and  $\rho_{TT}$  (right) and their commonly selected threshold value (vertical dashed black line). Thick grey lines indicate the average slope of curves



of advected turbulent kinetic energy with the production term along a given trajectory is at the basis of the proposed criterion. A robust implementation of the method applies the median operator for the production term and an additive term that avoids a null denominator.

The method is verified with data from three different experiments, including planar, tomographic 3D-PTV and 3D-PTV measurements by coaxial volumetric velocimetry.

The outlier detection principle based on turbulence transport detects the edges of regions of outliers, mostly caused by extensive regions of light reflection or shadows in the measurements. The method compares favourably with the state-of-the-art UOD from the analysis of the experimental

datasets. Considering the two optimal thresholds,  $r_0^* = 2$  and  $\rho_{TT} = 5$ , the latter yields a significantly higher detection ratio  $\eta_d$  with a comparable and small amount of false positives ( $\eta_{fp}$ ). For extended regions of outliers, the method requires multiple erosion iterations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated

otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Adatrao S, Bertone M, Sciacchitano A (2021) Multi- $\Delta t$  approach for peak-locking error correction and uncertainty quantification in PIV. *Meas Sci Technol* 32:054003
- Agüera N, Cañero G, Astarita T, Discetti S (2016) Ensemble 3D PTV for high resolution turbulent statistics. *Meas Sci Technol* 27:12
- Ahmed SR (1984) Influence of base slant on the wake structure and drag of road vehicles. *J Fluids Eng* 105(4):429–434
- Azizli I, Dwight RP (2015) Solenoidal filtering of volumetric velocity measurements using Gaussian process regression. *Exp Fluids* 56:198
- Duncan J, Dabiri D, Hove J, Gharib M (2010) Universal outlier detection for particle image velocimetry (PIV) and particle tracking velocimetry (PTV) data. *Meas Sci Technol* 21:057002
- Faleiros DE, Tuinstra M, Sciacchitano A, Scarano F (2019) Generation and control of helium-filled soap bubbles for PIV. *Exp Fluids* 60:40
- Godbersen P, Schröder P (2020) Functional binning: improving convergence of Eulerian statistics from Lagrangian particle tracking. *Meas Sci Technol* 31:095304
- Hart DP (2000) PIV error correction. *Exp Fluids* 29:13–22
- Higham JE, Brevis W, Keylock CJ (2016) A rapid non-iterative proper orthogonal decomposition based outlier detection and correction for PIV data. *Meas Sci Technol* 27:125303
- Hinze J (1975) *Turbulence*, 2nd edn. McGraw-Hill, New York
- Ikhennicheu M, Druault P, Gaurier B, Germain G (2020) Turbulent kinetic energy budget in a wall-mounted cylinder wake using PIV measurements. *Ocean Eng* 210:107582
- Jux C, Sciacchitano A, Schneiders JFG, Scarano F (2018) Robotic volumetric PIV of a full-scale cyclist. *Exp Fluids* 59:74
- Lazar E, DeBlauw B, Glumac N, Dutton C, Elliott G (2010) A practical approach to PIV uncertainty analysis. *AIAA Paper* 4355:30
- Masullo A, Theunissen R (2016) Adaptive vector validation in image velocimetry to minimise the influence of outlier clusters. *Exp Fluids* 57:33
- Raffel M, Willert CE, Scarano F, Kähler CJ, Wereley ST, Kompenhans J (2018) *Particle image velocimetry: a practical guide*, 3rd edn. Springer, Berlin
- Raiola M, Discetti S, Ianiro A (2015) On PIV random error minimization with optimal POD-based low order reconstruction. *Exp Fluids* 56:75
- Saredi E, Sciacchitano A, Scarano F (2020) Multi- $\Delta t$  3D-PTV based on Reynolds decomposition. *Meas Sci Technol* 31:084005
- Scarano F, Ghaemi S, Caridi GCA, Bosbach J, Dierksheide U, Sciacchitano A (2015) On the use of helium-filled soap bubbles for large-scale tomographic PIV in wind tunnel experiments. *Exp Fluids* 56:42
- Schanz D, Gesemann S, Schröder A (2016) Shake-The-Box: Lagrangian particle tracking at high particle image densities. *Exp Fluids* 57:70
- Schneiders JFG, Caridi GCA, Sciacchitano A, Scarano F (2016) Large-scale volumetric pressure from tomographic PTV with HFSB tracers. *Exp Fluids* 57:164
- Schneiders JFG, Scarano F, Jux C, Sciacchitano A (2018) Coaxial volumetric velocimetry. *Meas Sci Technol* 29:065201
- Sciacchitano A (2019) Uncertainty quantification in particle image velocimetry. *Meas Sci Technol* 30:092001
- Song X, Yamamoto F, Iguchi M, Murai Y (1999) A new tracking algorithm of PIV and removal of spurious vectors using Delaunay tessellation. *Exp Fluids* 26:371–380
- Stanislas M, Okamoto K, Kähler CJ, Westerweel J (2005) Main results of the Second International PIV Challenge. *Exp Fluids* 39:170–191
- Wang HP, Gao Q, Feng LH, Wei RJ, Wang JJ (2015) Proper orthogonal decomposition based outlier correction for PIV data. *Exp Fluids* 56:43
- Westerweel J (1994) Efficient detection of spurious vectors in particle image velocimetry data. *Exp Fluids* 16:236–247
- Westerweel J, Scarano F (2005) Universal outlier detection for PIV data. *Exp Fluids* 39:1096–1100

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.