A probabilistic particle tracking framework for high particle densities

D. Schiepel[1,*], S. Herzog[2], R. Barta[1], C. Wagner[1,3]

1: Institute for Aerodynamics and Flow Technology, German Aerospace Center, Germany
2: Third Institute of Physics - Biophysics, University of Göttingen, Germany
3: Institute for Thermodynamics and Fluid Mechanics, Technische Universität Ilmenau, Germany
*Corresponding author: Daniel.Schiepel@dlr.de

Keywords: Particle Tracking Velocimetry, Gaussian Mixture Model, Rayleigh-Bénard convection

ABSTRACT

A framework for particle tracking velocimetry at high particle densities (HD-PTV) based on a Gaussian Mixture Model (GMM) is presented. This new approach is validated by tracking synthetic particles generated for a generalized turbulent pipe flow defining the ground truth. For a step size per time step of $\delta_S = 14\,\mathrm{px}$ and a particles per pixel (ppp) density of 0.09 the framework tracks about 90% of the ground truth particles (percentage of matched particles, pmp) already after 9 time steps without generating any ghost particles. For a lower step size of $\delta_S = 7\,\mathrm{px}$, corresponding to a higher temporal resolution of the flow, and the lowest investigated particle density ppp $= 0.02$ a constant pmp close to 100% is reported. A decrease on pmp to 80% is found for the highest ppp $= 0,11$ - corresponding to about 45000 particles in total. Increasing the step size per time step to $\delta_S = 14\,\mathrm{px}$ results in a similar sloping curve and pmp that are generally 5% lower compared to the lower step size.
The approach is further successfully applied to a well-known experimental tracking problem, i.e. particle tracking in turbulent Rayleigh-Bénard convection, for which the motion of about 28500 particles is tracked. With track lengths up to 250 times steps the occuring structures and velocities are investigated and agree well with previous studies based on tomographic particle image velocimetry using the same data.
Thus, it is concluded that the presented HD-PTV framework is an appropriate tool for the flow analysis even at high particle densities.

1. Introduction

The presented framework for particle tracking velocimetry at high particle densities (HD-PTV) is based on a Gaussian Mixture Model (GMM). It is an extension of the state-of-the-art iterative reconstruction of individual particles by a continuous modeling of the particle trajectories considering the position and velocity as coupled quantities. Generally speaking, PTV methods rely on two common steps: First, 3D particle reconstruction where spatial information is extracted from the camera images and second, the linking of matching particle using another set of criteria to form tracks.

The current state-of-the-art framework for time-resolved 3D particle tracking velocimetry is the commercially available *Shake-The-Box* (STB) software, where particle positions in subsequent images are predicted by extrapolating trajectories with a third-order polynomial, as described in Schanz et al. (2016). A prerequisite for this is the combination of calibration methods like volume self-calibration, see Wieneke (2008), with iterative particle reconstruction (IPR), described in Wieneke (2013), and image matching by a so-called shaking step, see Schanz et al. (2016).

In contrast the here presented approach is based on a GMM, see Reynolds (2015) and Garcia et al. (2009), for the trajectory modeling to describe and predict the movement of the particles. The main advantage of the presented HD-PTV is: (1) modeling of the trajectories as coupled dimensions (i.e.: position and velocity) resulting in a high prediction accuracy, (2) the ability to change the start/end point without a total recalculation of the entire track allowing high computational efficiency, (3) creating a compact representation of the determined tracks requiring less memory space and (4) the ability to model periodic trends in the particle motion.

The here presented particle tracking velocimetry framework consists of an image pre-processing, 3D particle reconstruction using triangulation, an initialization step to estimate the velocities, a particle movement prediction using probabilistic motion prediction described by the introduced GMM and an iterative particle reconstruction for the later time steps. Here, we focus on the presentation of the probabilistic motion prediction. The entire framework in the context of PTV and the accuracy assessment of the method using the ground truth information of synthetic data set of a generalized pipe flow is presented. More over, the real-world functionality is demonstrated by applying the method to an experimental case of turbulent Rayleigh-Bénard-convection in a cubic sample.

2. Particle Trajectory Modeling

Using a standard particle detection and triangulation the particles P with position $\vec{X} \in \mathbb{R}^3$ are first reconstructed by the framework, details can be found in Herzog et al. (2021). For the modeling of the trajectories of P in time t, three (one for each spatial dimension) GMMs, see Reynolds (2015) and Herzog et al. (2021), are used. The aim of this approximation is to determine a function mapping based on the currently observed history for predicting the most likely particle motion.

However, since P may perform a nonlinear movement in a turbulent flow, as already stated in Aref (1990), predicting the particle position $\vec{X}$ is not a trivial task, see Muthu & Murali (2021). In order to deal with it, the GMM modeling is used instead of a spline for the extrapolation used in Wieneke (2013). More precisely, the employed GMM can model both the position $\vec{X}_0, \vec{X}_1, \ldots, \vec{X}_{t-1}$ and velocity $\dot{\vec{X}}_0, \dot{\vec{X}}q_1, \ldots, \dot{\vec{X}}_{t-1}$ of the particle. The density of a Gaussian mixture model, as described in Reynolds (2015), is given by

$$p(x) = \sum_{k=1}^{K} \pi_k \mathcal{N}(x|\mu_k, \Sigma_k), \tag{1}$$

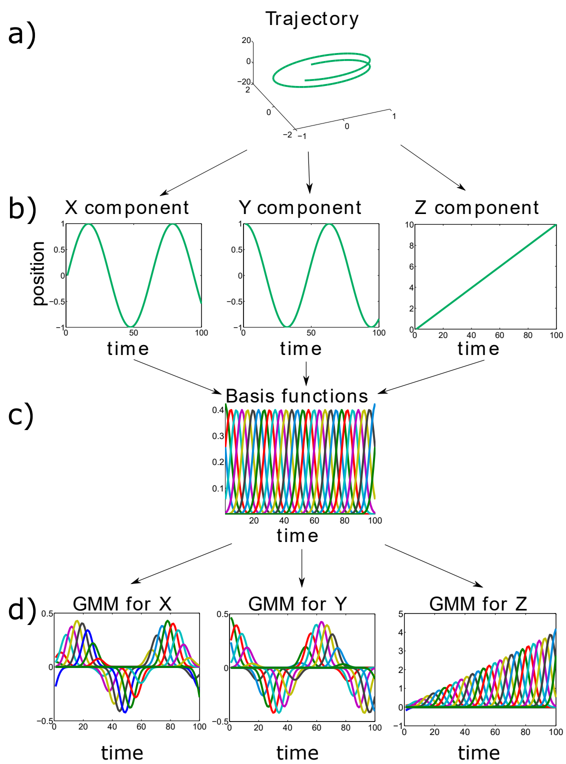

Figure 1. Representation of a three-dimensional trajectory with three GMMs. The first line a) shows an example of a trajectory and the second line b) the X, Y and Z position components of the trajectory. c) represents the set of basis functions from Eq. (1). By using the basis functions and the components from b), w is calculated. By scaling the basis functions with the obtained w the basis functions preserve the shape of the desired component. The result is illustrated in d).

where x is the d-dimensional random variable, $\mathcal{N}(x|\mu_k m, \Sigma_k)$ is a multivariate normal distribution, with mean μ_k and a covariance matrix Σ_k. To derive the parameters for equation (1), the Expectation-Maximization (EM) algorithm Moon (1996) is used. As a model linear basis functions with weights $w \in \mathbb{R}^K$ are used. Here, the basis functions for the position are defined by the product of normal distributions and accordingly the basis functions for the velocity are their time derivatives. Using this representation only the weights w need to be stored for each trajectory. This process is illustrated in Figure 1. The first line a) shows an example of a generic trajectory and the second line b) the X, Y and Z position components of the trajectory. c) represents the set of basis functions from equation (1). By using the basis functions and the components from b), w is calculated. By scaling the basis functions with the obtained w the basis functions preserve the shape of the desired component. The result is illustrated in d).

3. Synthetic pipe flow

In order to validate and benchmark our framework for the tracking of particles at high densities, a synthetic case of a generalized pipe flow is set up. The simulated flow is used as a ground truth case to directly compare the reconstruction results based on the number and accuracy of the generated trajectories. The synthetic case is designed to evaluate the capability of the HD-PTV framework to fulfill five requirements related to: three-dimensional particle movement, temporal resolution, particles leaving the domain, contradicting movement of the particle images between the time steps and varying particles per pixel (ppp) densities.

For the synthetic case of a generalized pressure-driven turbulent flow through a pipe, N_p particles are initially distributed within a cubic volume with dimensions of $W \times H \times D = 500 \times 500 \times 500\,\text{mm}^3$. Each particle P has the attributes position $\vec{X} \in \mathbb{R}^3$ and the intensity $I \in \mathbb{R}^{n_c}$ for the individual intensities on the n_c camera images. In the considered cases $n_c = 4$ virtual cameras observe the particle distributions within the 3D cubic volume centered around $(0,\ 0,\ 0)$ corresponding to boundaries $X_{\text{max/min}} = \pm(250,\ 250,\ 250)$.

In Figure 2 an exemplary particle distribution consisting of an outer (blue) and an inner structure (green) is illustrated for a total particle number of 28000 and a step width of $\delta_S = 7\,\text{px}$. The front view in Figure 2a) shows the annular structures with different radii and the rotational movement around the axis by accordingly colored arrow. The particle motion in longitudinal direction is illustrated by the side view in Figure 2b) and the dominant movement direction is indicated by arrows.

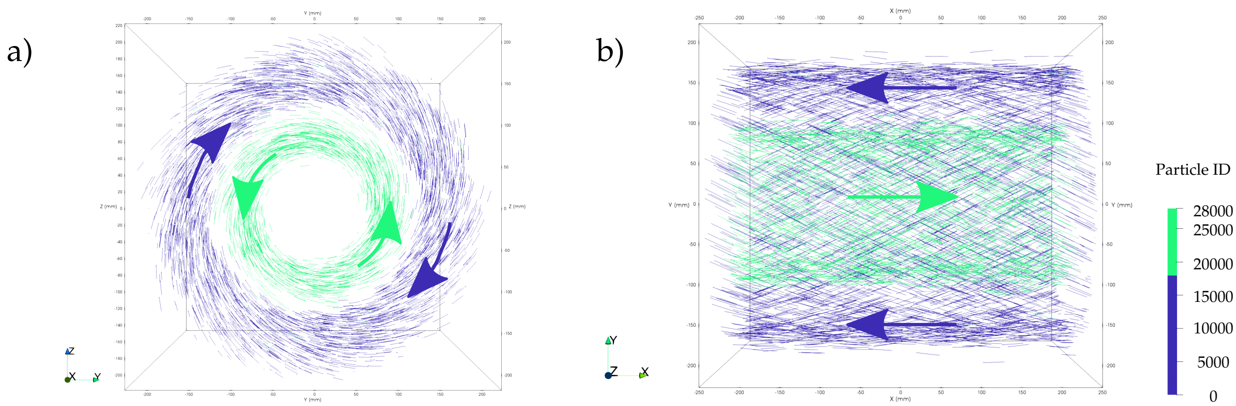

Figure 2. Particle distribution organized in two counter-rotating aligned annular structures with opposite mean flow directions presented in blue and in green. In a) the perspective is chosen along the X axis to illustrate the rotational movement around the axis which is also indicated by accordingly colored arrows. In b) the view along the Z-axis reflects the movement in X direction with the dominant movement direction indicated by arrows.

3.1. Performance Assessment

The proposed particle tracking approach is used to track the particle motion in comparison to the ground truth of the particle trajectories generated in the above-defined synthetic test case. To evaluate the performance of the method, the percentage of matched particles (pmp) is introduced and analyzed for particle densities between $0.015\,\text{ppp}$ and $0.107\,\text{ppp}$. Two step widths given in pixels per time step, $\delta_s = 7$ and $14\,\text{px}$ representing the mean particle velocities, and thus, the temporal resolution, are investigated. Each reconstructed and true particle is identified by an unique ID. The earlier is classified as matched, if it is within $1.5\,\text{mm}$ of a true particle and both particle IDs persist over time. The pmp is then defined by the number of matched particles compared to the true total particle number and a double assignment of a true particle is not permitted.

A basic requirement for the method is the reconstruction of large flow structures. Figure 3 shows the reconstructed structures for the synthetic data set at time step 50, where parts of the reconstructed particle trajectories extending over 10 time steps are depicted. They are color-coded by the starting time step of the particle trajectory. The comparison with the synthetically generated structures shown in Figure 2 reveals the qualitative agreement. Further, most of the trajectories are colored in yellow and started accordingly at a time step close to $t = 0$. This proves the capability

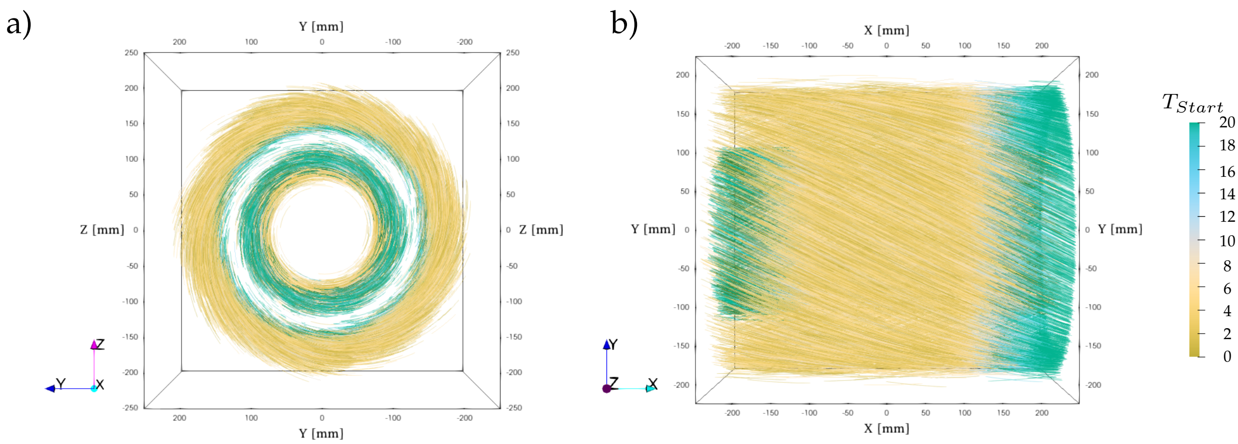

Figure 3. Reconstruction of the aligned annular structures from the synthetic data set at time step 50. Particle trajectories reflecting 20 time steps are depicted and color-coded with the starting time step of the particle trajectory.

of the track initialization and maintaining long trajectories. Only the two regions for the newly entering particles in Figure 3b) (right for outer structure and left for the inner one) have higher starting time steps up to $t = 50$ reflected by the green color coding. This demonstrates the ability of adding new particles and trajectories.

It is well-known for PTV methods that a certain number of time steps is required to accumulate all particles for the tracking system in the first processing step. Thus, an important quality criterion for those methods is the number of required time steps to acquire the maximum possible pmp. Figure 4a) shows the pmp and the number of matched particles obtained for 28000 true particles as a function of the time step. A close-up view of the first 20 time steps in the lower right corner is presented. After 9 time steps a plateau of about 90% matched particles corresponding to a total number of 25200 particles is reached. It should be noted that the method does not reconstruct ghost particles. Additionally, a significant length of the reconstructed tracks is necessary to obtain a sufficient amount of connected information for the paths of the particles. Figure 4b) shows the frequency distribution of the track length measured in time steps. For short track lengths of 0 to 70 time steps a low frequency of $1.0 \cdot 10^3$ is reported with a large increase to $9.0 \cdot 10^3$ for longer tracks between 100 and 150 time steps respectively. Considering an average velocity of $4.0\,\mathrm{mm/s}$ and a sample length of 500, a maximum of 125 time steps can be reached on average before the particles leave the domain. Figure 4c) shows a violin plot of the three velocity components. The colored area is normalized thus showing the relative distribution per component. The X component exhibits two separated agglomerations around $\pm 4\,\mathrm{mm/s}$. The contribution at the negative values is larger compared to the ones. This results from the higher particle number in the outer structure with negative movement direction in X. The distributions for the Y and Z velocity components are similar with two agglomerations at the higher velocities. This is due to the circular motion coupling both components: When the velocity, is high in Y, it is low in Z and vice versa. Both the qualitative behaviour and quantitative distribution comply with the ground truth.

Figure 5 reflects the matched particles and pmp as a function of the true particle number to demon-

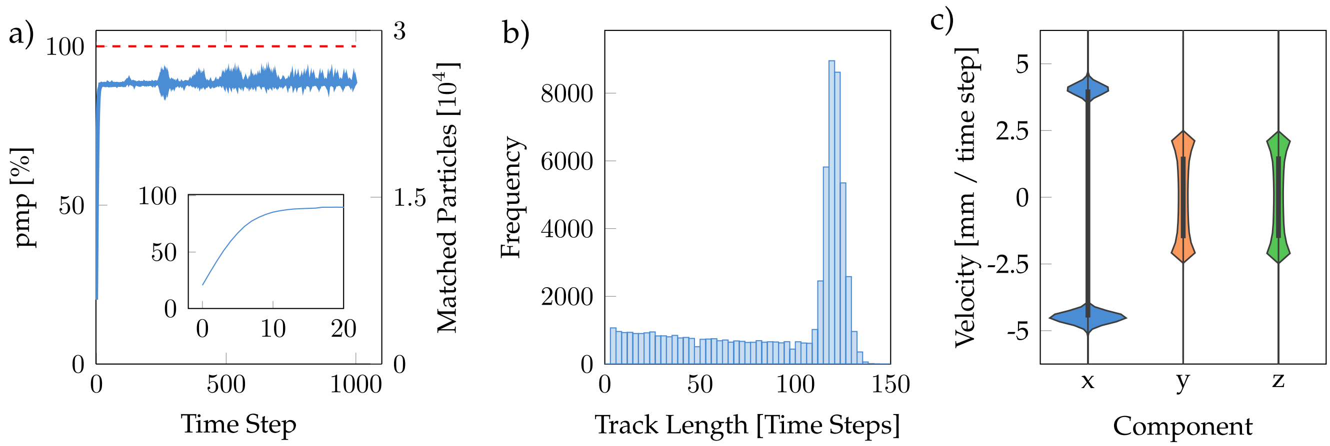

Figure 4. Statistics for the synthetic data set containing 28000 true particles. a) shows the number of matched tracks and pmp as a function of the time step in blue. The true particles number is indicated as a red line. b) shows the frequency distribution of the track length. c) shows the violin plot of the velocities of the tracked particles, where the estimated probability density is given by the colored shape, the interquartile range by the thicker black line and the whole data range by the thinner line.

strate the performance of the framework. The curve starts close to $100\,\%$ pmp for the lowest particle number. For an increasing ppp, a continuous decrease of the pmp is observed. For ppp = 0.05, the pmp is about $92.5\,\%$. Finally, for the highest ppp of 0.11, a pmp of $80\,\%$ is achieved. An increased particle velocity is investigated represented by the step width per time step of $\delta_S = 14\,$px reflected by the orange line in Figure 5. The general behavior of this reconstruction efficiency is similar to the previous case. For the lowest particle density of 0.005 ppp an efficiency of $97\,\%$ pmp is reported with a steady decrease to $77\,\%$ pmp for a high ppp of 0.1. While for the lowest particle number, the difference between the two particle velocities is about $1\,\%$ pmp, it increases to about $2.5\,\%$ for the highest ppp. The general decrease of the pmp is not surprising, since the association of particles with stable tracks becomes more ambiguous for increasing ppp. Still, while pmp = 99% for ppp= 0.007 corresponds to stable information from 7000 particle tracks, a significant increase in information density is achieved with 34000 stable tracks for the highest ppp.

The results were obtained on a workstation computer featuring an Intel Xeon Platinum 8268 with 24 cores (48 threads) and $256\,$GB RAM. On average (over all cases) about 2981 seconds were needed to reconstruct one time step, when running on a single thread and 73.7 seconds when running on 48 threads. It should be noted, however, that the execution time depends heavily on the ratio of successfully extended trajectories and newly triangulated particles. While the extension of n trajectories increases the execution time almost linearly, the execution time increases according to a cubic law when n new particles have to be triangulated and assembled into trajectories. Especially in the cases with high particle densities, the execution time per time step can differ by orders of magnitude. During the parallel execution of the code $8\,$GB to $12.3\,$GB of RAM was allocated.

In general the proposed method is applicable for the investigated parameter range of ppp ≤ 0.100 and $\delta_S \leq 14\,$px.

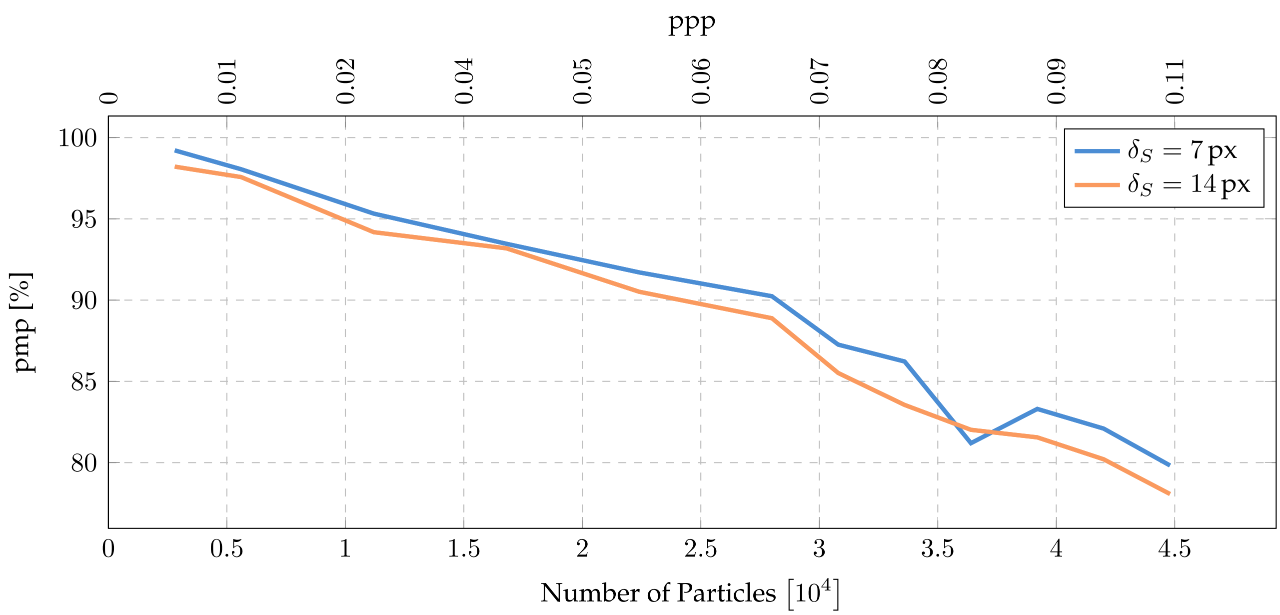

Figure 5. Number of matched particles and pmp as a function of the true particle number for two $\delta_S = 7$ and 14 px.

4. Turbulent Rayleigh-Bénard convection

The following section presents the applicability to experimental data by investigating PTV data acquired in a cubic Rayleigh-Bénard convection (RBC) sample with side length of $l = 500$ mm filled with water. The sample is illuminated by a LED array and the particle images are recorded using four PCO Pixelfly CCD cameras operating at 5 Hz. For the here considered case the experimental settings are $T_C = 18\,°C$, $T_H = 24\,°C$, with an average sample temperature of $\overline{T} = 21\,°C$ and $\Delta T = 6$ K. The corresponding characteristic numbers are the Prandtl number of

$$Pr = \frac{\nu}{\kappa} = 6.9$$

and the Rayleigh number of

$$Ra = \alpha\, g\, l^3\, \frac{\Delta T}{\nu \kappa} = 1.0 \cdot 10^{10}$$

with the kinematic viscosity ν the thermal diffusivity κ, the thermal expansion coefficient α and the gravitational acceleration g. A full description of the experimental apparatus and setup can be found in Schiepel et al. (2013).

Here, the applicability of the framework is proven by the obtained particle trajectories presented in Figure 6. Statistics are acquired by evaluating 250 time steps. For time step $t = 100$, the reconstructed particles are presented as small spheres with the particle path for the previous 50 time steps attached as a tail. The latter is color-coded by the velocity magnitude. In Figure 6a) a view along the Z-axis is chosen to demonstrate the reconstruction of particle tracks in the entire volume. In addition, a large-scale flow structure filling the convection sample is visible. This was also found for the same data set by Schiepel et al. (2013) using tomographic PIV. New insight is gained by the reconstruction of a previously unresolved smaller corner flow reported at X and $Y \approx 50\,\text{mm}$. In Figure 6b) the perspective is rotated to a diagonal plane of the sample in order to reflect the orientation of the large-scale circulation (LSC) more clearly. The instantaneous velocity magnitudes range up to $20\,\text{mm/s}$ with the majority being significantly slower as reflected by the dominantly blue trajectories.

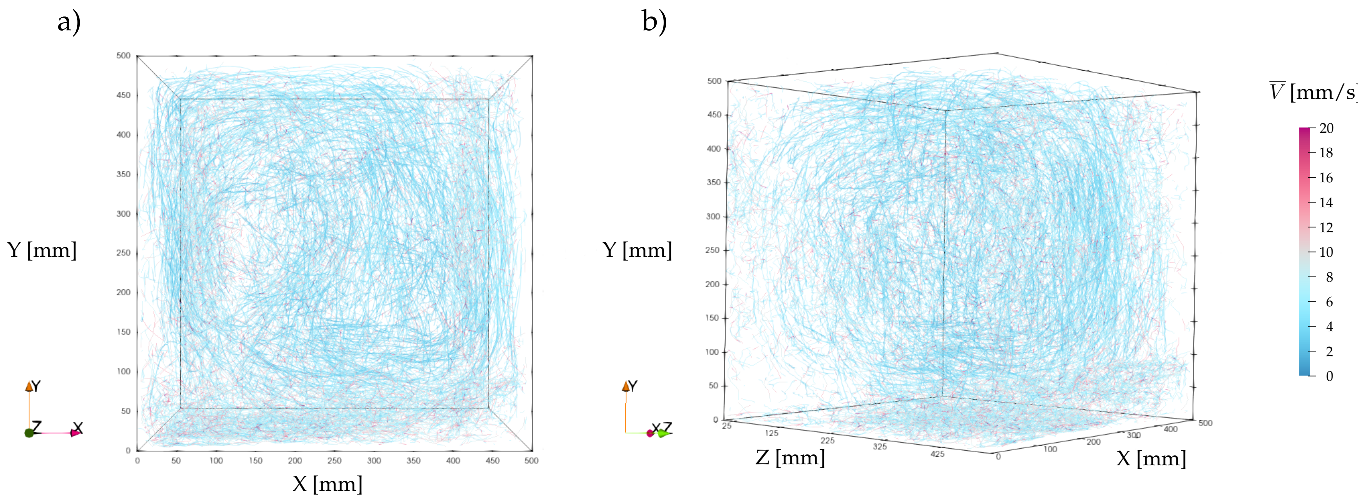

Figure 6. Reconstruction of the RBC data at time step 100. Particle trajectories reflecting 50 time steps are depicted and color-coded with the velocity magnitude.

In order to make quantitative statements on the flow occurring within RBC, the trajectories resulting from the tracking are further analyzed. The particle number for the enclosed RBC case is expected to be constant. This is confirmed in Figure 7 a). After only six time steps the number of simultaneously tracked particles is 27472 with a stable plateau of about 28500 particles reached at time step 10.

For PTV methods, typically long trajectories are desired. Figure 7b) reflects the frequency distribution of the track length. A distribution with a track length of up to 250 time steps and with the strongest contribution to smaller track length (≤ 50 time steps) is obtained. The evaluated data set contained 250 time steps thus determining the limit which indicates that the longest tracks reached the potential distance. While the synthetic case showed a clear indication at a maximum track length determined by the system, the experimental case is expected to show less pronounced dominant track length. The mean track length is 10.04 ± 10.03 and the median length is 7. Turbulent RBC is characterized by its velocity distribution and can therefore be used as a validation case for the presented HD-PTV framework. A violin plot of the velocities is presented in Figure 7c). Each velocity component shows a symmetric distribution. This is in agreement with the expectations in terms of a flow within an enclosed system driven by a dominant circular structure. The fact that the LSC is not exactly aligned with the diagonal plane explains the different spread in the frequency distribution of the velocity components.

For this experimental case, the execution time per time step was also recorded. The same Intel Xeon

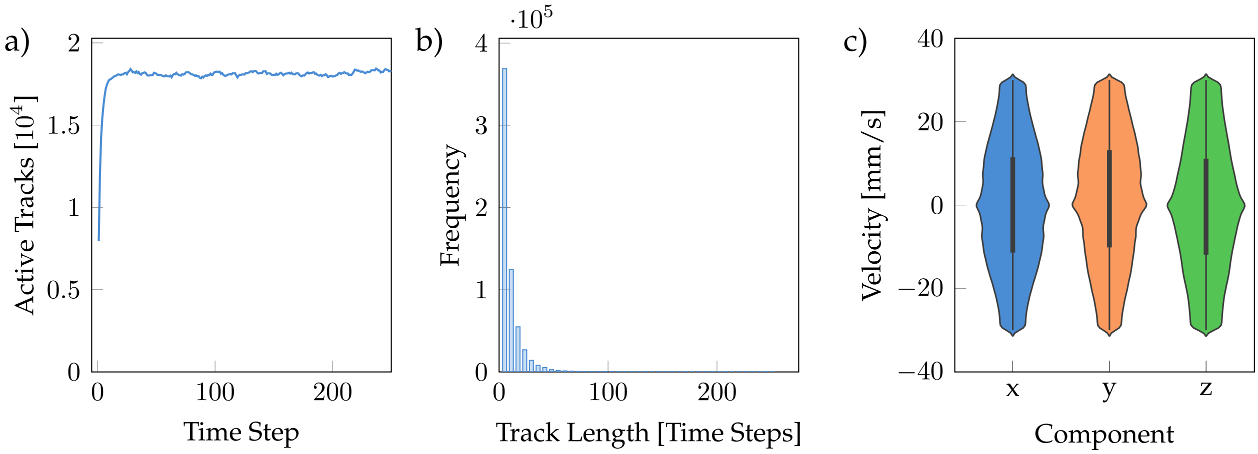

Figure 7. Same Figure structure as for Figure 4 but for the RBC case. a) shows the number of simultaneously tracked tracer particles in each time step. b) reflects the frequency distribution of the track length whereas c) shows a violin plot of the velocities.

Platinum 8268 achieved a average execution time of 2871 seconds per time step when running on one thread and 44 seconds when running with 48 threads. The comparatively higher execution time can be explained by the fact that much shorter trajectories are reconstructed for the experimental case than in the synthetic case. This implies that new particles have to be triangulated more often which is more costly. In total 10.5 GB RAM was allocated during the parallel execution.

The results from the experimental case justify the applicability of the developed framework for turbulent flows as it allows to reliably measure particle position, velocities and their associated trajectories.

5. Conclusion

A framework for particle tracking velocimetry at high particle densities (HD-PTV) based on a Gaussian Mixture Model (GMM) was presented. The approach is benchmarked by tracking particles generated in a synthetic generalized turbulent pipe flow which defines the ground truth. Considering this synthetic case, the approach returns about 90% matched particles after 9 time steps without generating any ghost particles. Further, for a step width per time step of $\delta_S = 7\,$px starting close to 100% pmp for the lowest particle density in terms of particles per pixels (ppp) of 0.02 a steady decrease to pmp $= 80\%$ is reported for the highest ppp $= 0.11$. Increasing the step width per time step to $\delta_S = 14\,$px results in a similarly declining curve and pmp values that are generally 5% lower. The presented results demonstrate that the introduced framework can reconstruct trajectories even under high particle densities. The employed GMM is designed to be as data-driven as possible, thus preventing the user from adding too much expert knowledge and thus potential biases.

Finally, the approach is successfully applied to a well-known experimental tracking problem: Particle tracking in turbulent Rayleigh-Bénard convection for which the motion of about 28500 particles is visualized using tracks of lengths up to 250 times steps and structures as well as velocities are in good agreement with previous studies on the same data.

Thus, the presented HD-PTV framework is an appropriate tool for the flow analysis even at high particle densities.

References

Aref, H. (1990). Chaotic advection of fluid particles. *Philosophical Transactions of the Royal Society of London. Series A: Physical and Engineering Sciences*, *333*(1631), 273–288.

Garcia, V., Nielsen, F., & Nock, R. (2009). Levels of details for gaussian mixture models. In *Asian conference on computer vision* (pp. 514–525).

Herzog, S., Schiepel, D., Guido, I., Barta, R., & Wagner, C. (2021, Oct 17). A Probabilistic Particle Tracking Framework for Guided and Brownian Motion Systems with High Particle Densities. *SN Computer Science*, *2*, 485-505.

Moon, T. K. (1996). The expectation-maximization algorithm. *IEEE Signal processing magazine*, *13*(6), 47–60.

Muthu, J. S., & Murali, P. (2021). Review of chaos detection techniques performed on chaotic maps and systems in image encryption. *SN Computer Science*, *2*(5), 1–24.

Reynolds, D. (2015). Gaussian mixture models. *Encyclopedia of biometrics*, 827–832.

Schanz, D., Gesemann, S., & Schröder, A. (2016). Shake-the-box: Lagrangian particle tracking at high particle image densities. *Exp. Fluids*, *57*(5), 70.

Schiepel, D., Bosbach, J., & Wagner, C. (2013). Tomographic Particle Image Velocimetry of Turbulent Rayleigh-Bénard Convection in a Cubic Sample. *Journal of Flow Visualization and Image Processing*, *20*(1-2), 3–23.

Wieneke, B. (2008). Volume self-calibration for 3D particle image velocimetry. *Exp. Fluids*, *45*(4), 549–556.

Wieneke, B. (2013). Iterative reconstruction of volumetric particle distribution. *Meas. Sci. Technol.*, *24*, 024008.