REGULAR PAPER

Moritz Heinemann ⬤ · **Steffen Frey** ⬤ · **Gleb Tkachev** ⬤ · **Alexander Straub** ⬤ · **Filip Sadlo** ⬤ · **Thomas Ertl** ⬤

Visual analysis of droplet dynamics in large-scale multiphase spray simulations

Abstract We present a data-driven visual analysis approach for the in-depth exploration of large numbers of droplets. Understanding droplet dynamics in sprays is of interest across many scientific fields for both simulation scientists and engineers. In this paper, we analyze large-scale direct numerical simulation datasets of the two-phase flow of non-Newtonian jets. Our interactive visual analysis approach comprises various dedicated exploration modalities that are supplemented by directly linking to ParaView. This hybrid setup supports a detailed investigation of droplets, both in the spatial domain and in terms of physical quantities . Considering a large variety of extracted physical quantities for each droplet enables investigating different aspects of interest in our data. To get an overview of different types of characteristic behaviors, we cluster massive numbers of droplets to analyze different types of occurring behaviors via domain-specific pre-aggregation, as well as different methods and parameters. Extraordinary temporal patterns are of high interest, especially to investigate edge cases and detect potential simulation issues. For this, we use a neural network-based approach to predict the development of these physical quantities and identify irregularly advected droplets.

1 Introduction

Flow visualization has mainly been concerned with the analysis of single-phase flow, i.e., flows where a single type of fluid is involved (e.g., airflow around objects or liquid flow through machinery). However, in many science and engineering problems, two or even more phases are involved, such as water flowing in a domain containing air or in the dynamics of immiscible liquids. A major difficulty with the analysis of such multiphase flow is, however, its various degrees of complexity. On the one hand, it inherits all complexity of single-phase flow, whose visualization is subject to ongoing research. On the other hand, the dynamics and physics of the interface between the different phases are closely interrelated with the flow, as well as

M. Heinemann (✉) · G. Tkachev · A. Straub · T. Ertl
VISUS, University of Stuttgart, Allmandring 19, 70569 Stuttgart, Germany
E-mail: moritz.heinemann@visus.uni-stuttgart.de

S. Frey
Bernoulli Institute, University of Groningen, Nijenborgh 9 (Bernoulliborg), 9747 AG Groningen, The Netherlands

F. Sadlo
IWR, Heidelberg University, Im Neuenheimer Feld 205, 69120 Heidelberg, Germany
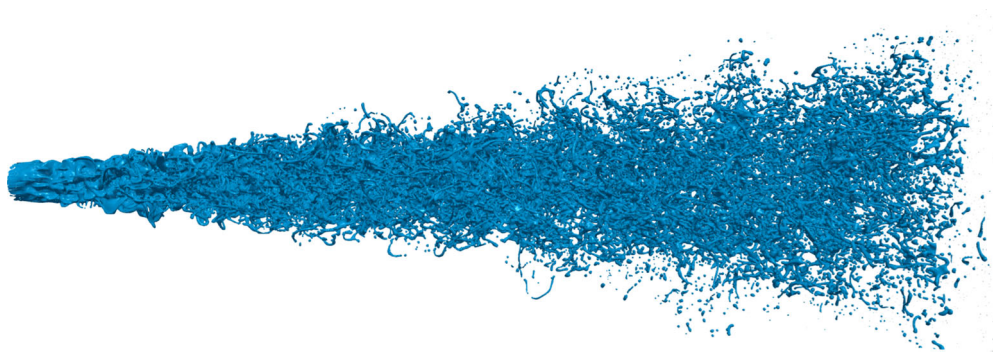
Fig. 1 Interface between liquid and air in the Jet Simulation dataset after the formation of a stable jet. Each time step features a rectilinear grid with a resolution of $1\,536 \times 512 \times 512$ cells covering a domain of 12 cm $\times$ 4 cm $\times$ 4 cm; 623 time steps represent a time span of 5.567 ms. The result is a volume of fluid field f and a velocity field $\mathbf{u}$. Both are given in cell-based representation

phenomena from solid mechanics such as collision, deformation, and adhesion. Furthermore, the volume of fluid method (VOF) (Hirt and Nichols 1981), which is typically used for simulating multiphase flow, further complicates analysis and interpretation. One reason is that the interface between the phases is not tracked but reconstructed at each time step during simulation, leading to inconsistencies between the flow in the vector field and the motion of the interface.

A phenomenon in two-phase flow with particularly high complexity is the formation of sprays. Sprays play an essential role in a wide range of natural phenomena and production, including precipitation, combustion, food processing, production and application of drugs, and cooling. In technical applications, sprays are typically generated by guiding the liquid through a spray nozzle, which produces an unstable jet that eventually breaks up into droplets. The dynamics of this breakup is highly complex, with primary breakup producing elongated components called ligaments, followed by a secondary breakup of ligaments into droplets.

All these processes and complexities make the resulting data extremely hard to navigate and analyze due to their spatiotemporal nature. While traditional visualization approaches are applicable to subsets and partial aspects of the data, they cannot provide an effective means for hypothesis forming and hypothesis testing for the entire data, in particular because the importance and interrelations of specific quantities and processes are buried in the discussed degrees of complexity. Another difficult aspect is that in order to be able to resolve the complex dynamics involved at small scales, high temporal and spatial resolution of the data is required. This results in data sizes quickly reaching terabytes, impeding the direct application of many types of advanced analysis procedures. While interactive exploration is required to analyze the complex, small-scale effects, this is challenging with the large data sizes and large numbers of droplets.

In this work, we aim to analyze the droplets in a two-phase DNS flow simulation of the breakup of a liquid jet in air (cf. Fig. 1). This simulation was conducted by Ertl (Ertl and Weigand 2017; Ertl 2019) using FS3D (Eisenschmidt et al. 2016) and considered water with 0.3% flocculant, leading to non-Newtonian fluid behavior. In this case, the simulation scientists' interest focused on the fully converged phase of the jet. Accordingly, we mainly consider a respective subsequence of 101 time steps covering 0.679 ms of the entire simulation (except for the collision and separation counters, which cover the complete 623 time steps). The total data size is around 7 TB (i.e., $\approx$1 TB for the focused subset of 101 time steps).

Our dedicated visual analysis approach provides different hypothesis forming tools to analyze droplet behavior in the data. Our goal is (i) to allow simulation and domain scientists to analyze what common types of droplets and their behaviors are, and (ii) to identify and study anomalous—i.e., uncommon—cases in depth. Below, after discussing related work (Sect. 2), we describe our workflow enabling the interactive visual analysis of large sets of droplets based on the extraction of meaningful quantities and advanced further automated analysis, including clustering and machine learning-based anomaly detection (Sect. 3). Our visual analytics system then provides different perspectives on abstract and spatial quantities and allows for detailed flow analysis with the original raw data (Sect. 4). Finally, we present insights gained from visual analysis (Sect. 5) before concluding (Sect. 6).

We present the first visual analysis approach for complex time-dependent multiphase flow data. By combining domain-specific clustering techniques with an artificial neural network for learning droplet

anomaly behavior, we developed a highly interactive framework to support flow scientists in the complicated tasks of analyzing terabytes of simulation data.

2 Related work

Our data stems from a CFD solver for multiphase flow simulation of incompressible fluids. This solver (Eisenschmidt et al. 2016) uses the VOF method (Hirt and Nichols 1981), in combination with Piecewise Linear Interface Calculation (PLIC) as interface reconstruction algorithm, which we also use in this work to analyze droplets (cf. Karch et al. (2013) for a discussion of PLIC in visualization).

The visual analysis of time-dependent flow fields is challenging. Aigner et al. (2007) discuss the consideration of time as an additional dimension in visual analytics. Bürger et al. (2007) integrate local feature detectors in the visual analysis of time-dependent flow simulations. Another line of work concerned itself with the interactive exploration of complex time-dependent flow simulations and real-world data (Doleisch et al. 2003, 2004a, b. Shi et al. (2009) introduced an approach to visually analyze time-dependent flow fields using pathlines in particular. An overview of different feature tracking techniques for flow visualization was presented by Post et al. (2003). Further Theisel and Seidel (2003) show a streamline-integration-based method for feature tracking in instationary vector fields. Garth et al. (2004) discuss an approach for tracking singularities within a vector field. In contrast to these previous works, while we also employ techniques from flow visualization for the detailed investigation of individual droplets, our specific focus on large numbers of droplet data features unique challenges that we aim to address in this work. Further previous work already focused on feature tracking in the context of large-scale datasets (Dutta and Shen 2016) and calculating tracking information in situ (Biswas et al. 2020).

Combining flow visualization and machine learning, Tzeng and Ma (2005) used neural networks to generate adaptive transfer functions based on user input. Bai et al. (2017) applied linear discriminant analysis to classify experimental images into actuated and unactuated flow. In this paper, we also apply neural networks to support flow visualization but use them to generate visual features that guide the analysis. Tkachev et al. (2019) trained neural networks on spatiotemporal volumes to detect irregular behavior. We use a similar idea for our anomaly detection in this work, but we apply our model to time series of extracted droplet quantities. Machine learning has also been applied in fluid simulations. Artificial neural networks are even used within solvers of the Navier–Stokes equation (Tompson et al. 2017) for accelerating the computation. In visualization in general, machine learning methods are particularly popular in visual analytics approaches (Endert et al. 2017).

This work started in the context of a master thesis (Heinemann 2018) which was reported in a very early version in a project report (Straub et al. 2019) and within a book chapter (Straub and Ertl 2020). The idea of using ML for prediction on droplet quantity time series is already mentioned there, but due to its overview character, no details of our technique were presented. Further, an early version of the surface view of our prototype was depicted there. This article introduces our framework and elaborates on new features like clustering, improvements to the 3D surface view, and extensions to the complete system, including the quantity and flow view and an analysis and discussion of the used dataset.

3 Preprocessing: extraction, clustering and anomaly detection

To make the data (interactively) explorable while still capturing the various degrees of complexity of droplet behavior, we first extract different quantities based on individual droplet instances (Sect. 3.1). Next, we establish temporal relationships by connecting individual droplet instances between the time steps to time-continuous droplet traces (Sect. 3.2). To get an overview of the droplet quantities, we use trace-based hierarchical clustering (Sect. 3.3). Also, we train a regression model to capture typical temporal patterns in droplets' quantities and then compute the deviations from this model to guide the researcher to anomalous cases in the sense of being uncommon. Akin to previous work (Tkachev et al. 2019), we chose artificial neural networks (ANNs) due to their generality, performance efficiency on large data (compared to, e.g., non-parametric models), and their successful applications across many diverse tasks (Sect. 3.4). Below, we mainly focus on applying and adapting the different methods involved to enable the interactive exploration of large numbers of droplets (Sect. 4). An overview of the order and dependencies of our processing steps can be seen in Fig. 2.
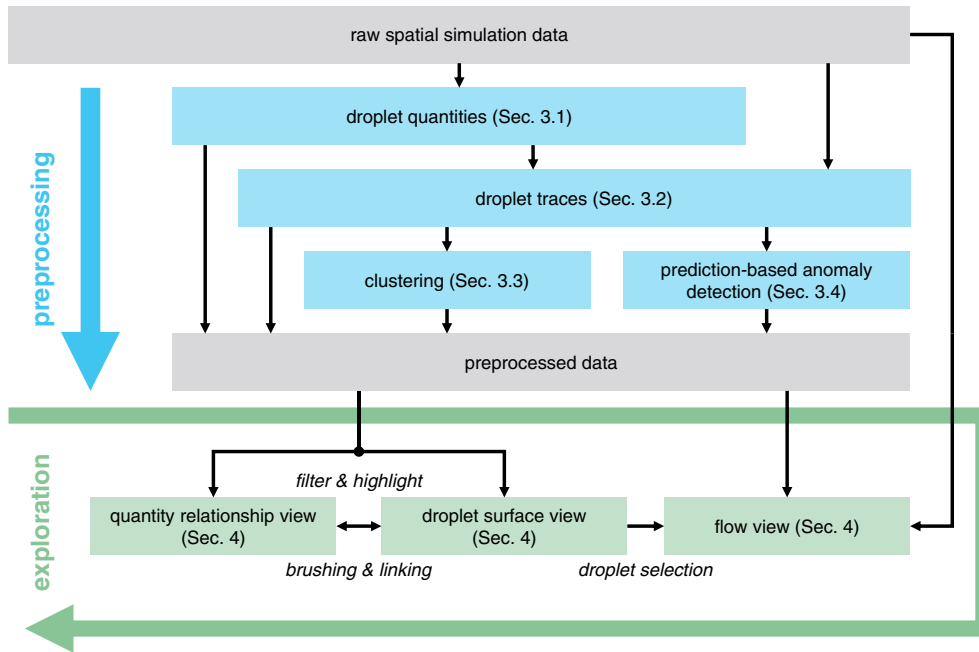
Fig. 2 Overview of the components of our visual analysis approach. This represents the execution order of our processing pipeline and shows data dependency of all steps described within Sect. 3 and Sect. 4. Each of the black arrows represents data usage from the previous step

3.1 Extraction of quantities

We extract different instantaneous quantities for each droplet instance, which capture the various degrees of complexity of droplet behavior. In collaboration with domain experts, the following set of instantaneous quantities has been determined, ranging from purely geometric to purely physical.

Data representation and segmentation. The first step consists in the identification of the individual droplets. The employed VOF approach in our two-phase flow simulation maintains a scalar field $f(\mathbf{x}) \in \mathbb{R}$ during the simulation. This field stores conceptually for each point $\mathbf{x} \in \mathbb{R}^3$ the volume fraction of the liquid, i.e., $f = 0$ representing only gas, $f = 1$ representing only liquid, and $0 < f < 1$ mixtures of both. Here, $f(\mathbf{x})$ is defined in a cell-based manner, i.e., the field stores this total fraction for each cell i individually, in piecewise constant representation. Thus, f_i represents the value of the f-field for cell i. Similarly, the velocity field $\mathbf{u}(\mathbf{x})$ is given in cell-based representation $\mathbf{u}_i$. In the domain, a droplet is commonly defined as the face-connected component of cells that exhibit $f_i > 0$. In practice, due to the numerical limitations of the simulation, we use the slightly modified definition $f_i > \tau_f$, with $\tau_f = 10^{-6}$. This value is defined by our domain experts, who run the simulation. Thus, we obtain the droplets by connected component labeling using region growing, resulting in a cell-based label field $l(\mathbf{x})$, where l_i stores the droplet identifier ι of the droplet that cell i belongs to.

A further peculiarity of two-phase flow simulation represents the interface between the two phases via PLIC (Youngs 1984). During the simulation, the interface is represented in a piecewise planar manner, i.e., within each cell with $0 < f < 1$, a planar patch represents the interface. It is determined by using the negative gradient of the f-field as a plane normal and choosing the position such that the volume below the patch is precisely the volume fraction represented by f_i. This piecewise linear representation per cell results in a discontinuous global interface containing gaps (see, e.g., Fig. 13). Using isosurface extraction for interface reconstruction would violate both the connected component definition of droplets and interface representation. Thus, it is dissuaded by domain experts. Therefore, we use the PLIC patches as geometric interface representation for both the computation of derived quantities and rendering, similar to Karch et al. (2013), at the cost of accepting the discontinuities.

Main instantaneous quantities. We want the extracted quantities to capture droplet behavior reasonably accurate, but at the same time, we are limited to quantities that can be calculated purely from the VOF and velocity field. Furthermore, we need a minimum of numeric stability in calculating them. While the overall

jet simulation is high resolution, single droplets within this dataset may only be resolved by few cells. Due to this, we have observed instabilities, mainly when calculating derivatives on the velocity fields. We therefore focus on derivative-free quantities and compute a total of 11 scalar quantities for each temporal instance of droplet ι: volume $\overline{\mu}_\iota$, area $\overline{A}_\iota$, area-to-volume ratio $\overline{\alpha}_\iota$, velocity $\|\overline{\mathbf{u}}_\iota\|$, momentum $\|\overline{\mathbf{p}}_\iota\|$, angular velocity $\|\overline{\omega}_\iota\|$, angular momentum $\|\overline{\mathbf{L}}_\iota\|$, total energy E_ι, kinetic energy $E_\iota^{\mathbf{u}}$, rotational energy E_ι^ω, and residual energy E_ι^δ. The formal definition of these quantities is provided within the appendix of this paper (Appendix A). Note that magnitudes are used in the case of vector-valued quantities for ensuring rotational invariance.

Additional quantities for evaluation and visualization. In addition to the main quantities used for the further preprocessing steps, we use additional quantities for visualization and evaluation. The polygons of the PLIC surface are stored to visualize droplet surfaces later. While we deemed derivation-based quantities to be not numerically stable enough to be used in automatic analysis, we still consider them useful when carefully employed in supplementing the analysis. To discern between droplets and ligaments, we compute spherical anisotropy measure from our segmented droplets. We achieve this by computing fractional anisotropy (Rosenberger et al. 2012; Basser and Pierpaoli 2011) of the covariance matrix of all cell centers comprising a droplet component ι, which we denote c_s, with $c_s = 0$ indicating linear or planar shape, and $c_s = 1$ for a perfectly spherical shape of a droplet. Furthermore, we compute the radius from the smallest surrounding sphere around the center of mass. While our main quantities above are designed to include droplet-local quantities, we include location-based quantities to refer to a position within the jet for analysis. Next to the absolute center of mass position, we use the distance of a droplet's center of mass to the jet's base axis and call it radius. In contrast, we do not include different velocity components, i.e., axial and radial velocity of a droplet, because the axial velocity is expected to approximate the overall velocity and the radial velocity is assumed to be a relatively small constant only depending on the position within the jet, as described by domain experts. The number of cells of droplet component ι is provided as a discretized alternative to $\overline{\mu}_\iota$ with two variants: (1) including the number of all cells contributing to a droplet and (2) the number of cells that are at least filled by 50% with the liquid phase. Additionally, we calculate the vortex core lines of each droplet according to Sujudi and Haimes (1995) and count the number of line segments to quantify the presence of a vortex. Furthermore, we save error flags during the computation process, i.e., if droplets are too small or the trace of a single droplet cannot be determined. These error flags have proven useful for filtering later on.

3.2 Trace generation

From the segmented droplet instances, we now establish a space-time graph depicting their temporal correspondence (i.e., a node is a droplet instance, and each edge a temporal relationship). Initially, the graph consists only of nodes and is extended by adding edges if we find correspondence between droplet instances in neighboring time steps. This is achieved by tracing imaginary particles from the center of each cell to the next time step (Karch et al. 2018). Unfortunately, using higher-order integration schemes for this, e.g., Runge–Kutta, would require interpolation in space and time, which is very likely to sample data from the gaseous phase and lead to erroneous results as pointed out by Karch et al. (2018). Therefore, we are using a forward Euler step:

$$\mathbf{x}(t + \Delta t) := \mathbf{x}_i(t) + \Delta t \mathbf{u}_i(t) . \tag{1}$$

An edge is added if $\mathbf{x}(t + \Delta t)$ belongs to a droplet, i.e., $\mathbf{x}(t + \Delta t)$ is located at time $t + \Delta t$ in a cell j with a valid droplet label. We also do a backward Euler step from the center $\mathbf{u}_k$ of each cell k that is part of a droplet at time $t + \Delta t$:

$$\mathbf{x}(t) := \mathbf{x}_k(t + \Delta t) - \Delta t \mathbf{u}_k(t + \Delta t) . \tag{2}$$

If $\mathbf{x}(t)$ belongs to a droplet at time t, i.e., if $\mathbf{x}(t)$ is located in a cell l with a valid droplet label at t, we add the respective edge (if not already present). If a node n_j has degree $d > 2$ at time t, and only one connected neighbor at time $t - \Delta t$, a breakup event has happened at node n_j at time t. If there is more than one connected neighbor at time $t - \Delta t$, coalescence is involved. Our analysis focuses on the dynamics of single droplets (not considering splits and merges; e.g., cf. Karch et al. (2018) for an analysis of these). Accordingly, we split the graph at nodes where more than two edges meet—i.e., coalescence and breakup events—, and base our approach on the isolated linear subgraphs, each representing the *trace* of a single

droplet over time. A trace contains droplet quantities as described above and serves as input to our regression model (Sect. 3.4).

3.3 Clustering

We want to cluster the droplets to reveal their different types of behavior in the simulation. With our distribution of droplet quantities, no meaningful results could be obtained with density-based clustering algorithms such as DBSCAN (Ester et al. 1996) (yielding just one or two major clusters and a large number of outliers). In contrast, hierarchical clustering according to Ward's method (Johnson 1967) has generated more expressive clusters in our experiments. However, hierarchical clustering exhibits quadratic memory complexity. The significant part is a distance matrix requiring $(n^2 - n)/2$ values to store. Handling our 1 000 000 droplet instances ($\approx 10\,000$ droplets over 100 time steps) would require at least 3.6 TB of memory and is therefore not possible on our machine. Instead, we aim to reduce the number of data points significantly. To achieve this, we limit ourselves for this modality to just one aggregate of each trace. Albeit at the cost of omitting temporal information, this can still serve the original purpose of identifying different droplet types (cf. Sect. 5). Note that our learning-based method presented below (Sect. 3.4) does not require this pre-aggregation.

3.4 Learning-based droplet anomaly detection

In our droplet anomaly approach, we first train a basic regression model on traces, using the surrogate task of predicting future values from a preceding time window. Assuming that the model does not overfit the training data (which we verify with a hold-out validation), it captures the most common and predictable patterns in droplet behavior. Then, we can quickly check droplets against the model to find ones that deviate from the typical behavior. In this work, we employ ANNs for their ability to handle a large number of elements and learn a useful data representation (Bengio et al. 2013).

As a surrogate task for training, we define a fixed-sized input window and slide it along the trace, applying the model at every window position to predict the next value in that trace. That is, given a trace of length n_t, and a fixed window size w, we obtain $n_t - w + 1$ *subtraces*, and for each subtrace, the model takes the first $w - 1$ time steps as input and predicts the droplet quantities at the last (w-th) time step in the subtrace. This time-window approach allows for comparably simple models to be used (lowering costs) and reduces the risk of overfitting because shorter subtraces are generally ambiguous and thus are more difficult to fit during the training process. We train a separate prediction model for each quantity to affirm that each quantity is given the same importance, avoiding compromises regarding accuracy as would be the case with multiple output variables. We split 20% of the data into a held-out validation dataset, using the rest for training. Of course, we want the training set to include as much data as possible, but the validation set needs to be a reasonably sized sample of the whole dataset. Due to containing a high amount of relatively similar droplets and a low amount of relatively widely spread outliers, a validation size smaller than 20% may introduce sampling bias. Beforehand, we normalize each quantity to have zero mean and a standard deviation of one, using the mean and standard deviation estimated on the training set. Each model is a fully-connected neural network with two hidden layers of 64 neurons using Rectified Linear Unit (ReLU) (Glorot et al. 2011) activation and a single linear unit in the output layer. The models are trained using the Adam
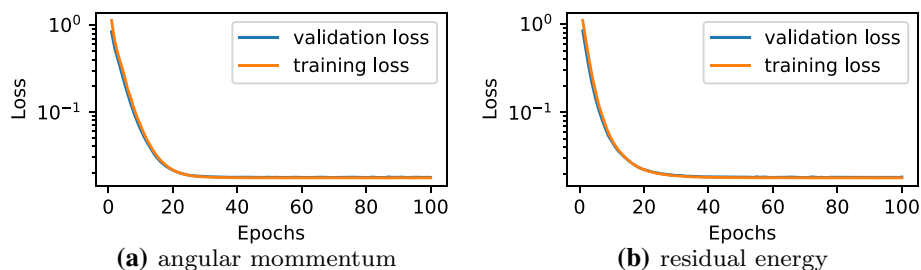

(a) angular mommentum   **(b)** residual energy

Fig. 3 An example of training and validation loss for the ANNs of angular momentum and residual energy. The curves for the remaining networks have a very similar shape

optimizer (Kingma and Ba 2014), with a learning rate 10^{-5}, and a batch size of 32 throughout 100 epochs until convergence of the validation MSE loss (Fig. 3). Our setup was chosen empirically by relying mostly on common values for most ML parameters, as our focus is more on the overall framework. We think there might be potential for further optimization, but we leave it for future work.

Once the models have been trained, we use them to perform predictions on all available subtraces. The first $w - 1$ time steps of a trace cannot be predicted by design and have to be omitted in the following. We then determine the difference between the actual and the predicted value of each quantity. In total, this results in an 11-dimensional vector of deviations, whose Euclidean norm finally yields total error $\bar{\varepsilon}$—our measure of *estimated droplet anomaly*.
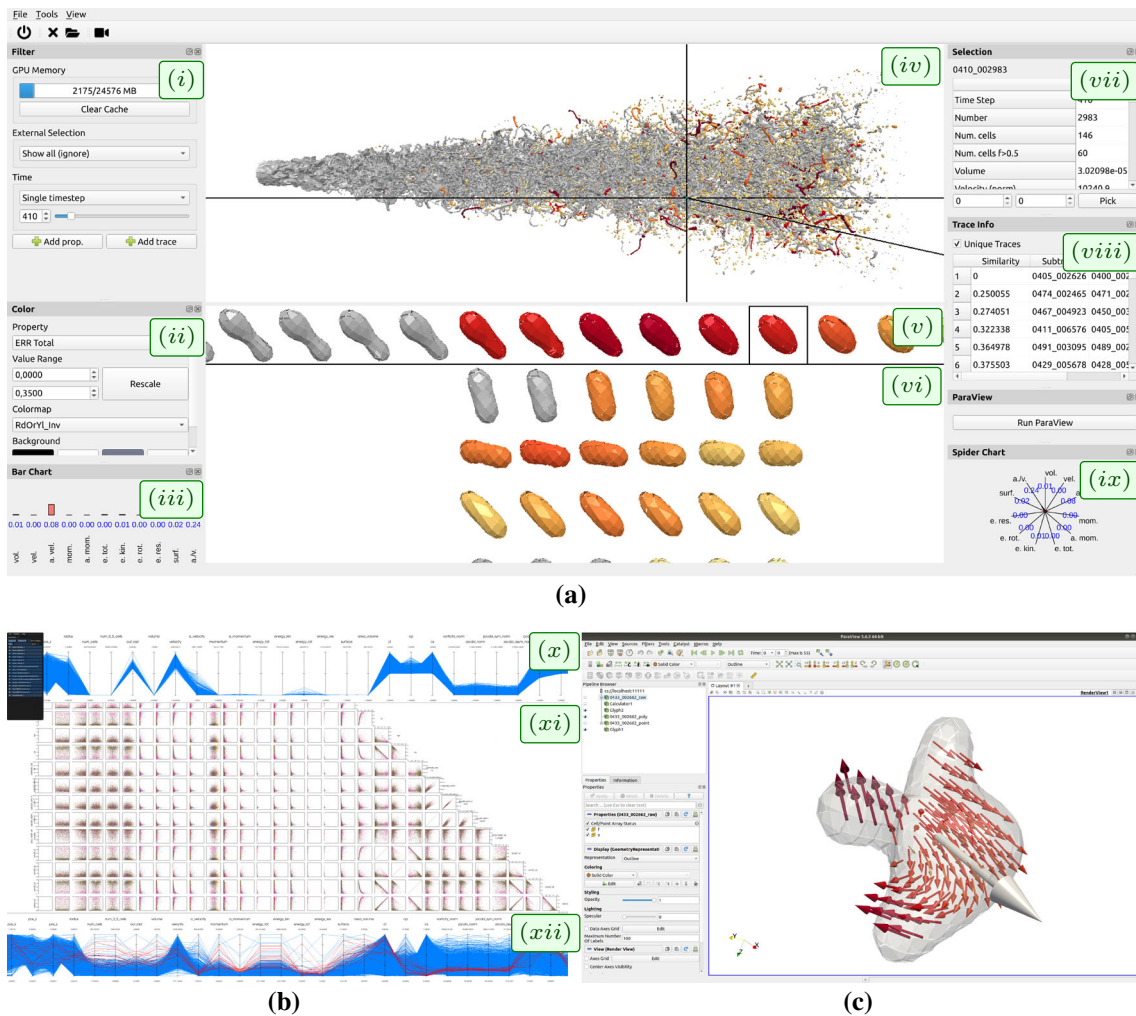

Fig. 4 System overview: **a** 3D surface view for showing droplet surfaces, consisting of filter panel (*i*), coloring tool (*ii*), bar chart (*iii*), main 3D view (*iv*) (here color mapped to anomaly measure $\bar{\varepsilon}$), time view (*v*) for selected droplet (crosshair in (*iv*)) with temporal scrolling. The similarity search (*vi*) provides cases similar to the selection, based on feature vector distance. In addition, we depict droplet instance information (*vii*), droplet trace similarities (*viii*) for (*vi*), and spider chart as a complement for (*iii*). **b** Quantity relation view: Input parallel coordinates plot for data filtering (*x*), scatterplot matrix (*xi*), and second parallel coordinates plot (*xii*) allows analysis of data. Moreover, any quantity can be mapped to color. **c** A ParaView instance is integrated within the system for analysis of the raw flow field for a single selected droplet and advanced flow feature extraction. A droplet selected in the 3D surface view can be automatically loaded into ParaView, including a useful default filter pipeline as shown within the figure

4 Visual analytics system

Our visual analysis system has three different views: the abstract quantity relationship view, the droplet surface view, and the flow view. These components and their interplay enable the detailed analysis of a large droplet collection by supporting various kinds of investigation. The droplet surface view displays a set of droplets in the 3D domain, allowing interactive exploration within the spatial and temporal context for selected droplet instances. The quantity relationship view provides an overview of extracted droplet characteristics and supports analyzing the interdependencies between them. Finally, the flow view can be used for the detailed investigation of a single selected droplet using complete raw input data. Different views are linked to supplement each other efficiently. An overview of the system can be seen in Fig. 4, and the different views are described in detail below.

Droplet surface view. The 3D droplet surface view allows a user to explore the droplet dataset spatially (Fig. 4a (*iv*)). Droplets of interest can be selected directly via picking (indicated with a crosshair). As showing all droplets would lead to significant occlusion and visual clutter, filtering the data is crucial (Fig. 4a (*i*)). We support filtering w.r.t. arbitrary physical or geometrical quantities, clusters, and prediction errors (and combinations thereof). Color mapping can flexibly depict chosen quantities, with the color-coded anomaly measure $\bar{\varepsilon}$ being the default choice (Fig. 4a (*ii*)). A user can navigate within the full trace of a selected droplet and explore its temporal evolution (Fig. 4a (*v*)). Below, similar traces are shown that were identified via feature vector distance, i.e., by Euclidean distance between droplet instances in the 11-dimensional quantity space. This helps, on the one hand, to assess the uniqueness of a droplet evolution, and on the other hand, the comparison to similar droplets can help to gain further insights into which commonalities or differences have led to certain behavior patterns (Fig. 4a (*vi*)). Additionally, we also provide the values for the selected droplet in our time view with a bar chart (Fig. 4a (*iii*)) and a spider chart (Fig. 4a (*ix*)), which can present its prediction error for different physical quantities. Detailed droplet instance information (Fig. 4a (*vii*)) and droplet trace similarities (Fig. 4a (*viii*)) are provided as well.

Quantity relationship view. The quantity relationship view is focused on displaying the different (abstract) droplet quantities in context with each other. Here, we use classical information visualization methods, namely parallel coordinates plots (PCP) and a scatterplot matrix (SPLOM), to present an overview of the physical quantities in the dataset as well as all derived quantities (e.g., clusters and prediction errors). In detail, this view consists of three separate components as shown in Fig. 4b. On top, we see the quantities of all droplets, shown within an interactive PCP (Fig. 4b (*x*)), where sliders on each axis can be used to filter the data. Positioned below are a scatterplot matrix (Fig. 4b (*xi*)) and a second PCP (Fig. 4b (*xii*)), in which we can analyze the filtered data. We provide both SPLOM and PCP to make use of the strengths of both visualization techniques. The PCP is ideal for getting an overview of the data and locating single data points with the overall value range context, while the SPLOM can show pairwise relations and is well-suited to identify patterns and relations within the data. Naturally, highlighting data in one view also will highlight the data in the other view. They are also linked to the droplet surface view (Fig. 4a), i.e., brushing within the SPLOM or PCP can be used to filter in the 3D surface view. This can help to obtain spatial context regarding location and surface shape to the abstract quantity data points. This component is implemented on top of MegaMol (Gralka et al. 2019), employing OpenGL to render millions of points and lines at highly interactive frame rates.

Flow view. For an in-depth exploration of the underlying flow field, e.g., to analyze the reason for a high anomaly, we further incorporate various classic flow visualization techniques by directly integrating ParaView (Ayachit 2015) into our system (Fig. 4c). ParaView is controlled from our application by loading the droplet data of the currently selected droplet within our 3D droplet surface view and automatically setting up the ParaView visualization pipeline. Not only the droplet itself is exported, but its full trace (in a droplet-local coordinate system for convenience), allowing animation. Moreover, we provide the precomputed surface and all other quantities. As frontend, the user has the classical fully functional ParaView interface.

5 Results

We will now analyze the Jet Simulation (cf. Sect. 1) in more detail to gain some insights and demonstrate our methods and system in practice. From the dataset, we extracted 1 043 168 droplet instances. 273 928 were omitted due to insufficient size (around two-thirds of them are artifacts at the simulation boundary), some more due to not being part of a trace with a minimum length of six. This results in 575 833 droplet

instances, which we use in our analysis. They are organized within 23 738 unique traces of sufficient length. These traces were split into 457 143 subtraces for training and validation data.

We used a machine featuring an Intel Core i7-8700K, 64 GB RAM, and an NVIDIA RTX Titan. The data had to be stored on an HDD due to its size. Therefore, disk I/O is the bottleneck for many of the tasks. The computation took ≈11 hours for segmentation (Sect. 3.1, CPU) and ≈25 hours for tracking (Sect. 3.2, CPU). Computation of the droplet quantities took ≈9.5 hours (CPU). Trace generation and additional data handling completed in ≈2 hours. The machine learning (Sect. 3.4, GPU) took ≈4 hours for the 11 ANNs, while clustering was finished in only ≈5 minutes (CPU). Our quantity view achieves more than 60 frames per second for a single time step and drops to around 5 frames per second when looking at the whole data at once (GPU).

5.1 Quantity relationships

We first consider a single time step with the quantity relationship view (Fig. 4b). As the jet we are looking at is fully converged, all time steps are quite similar regarding general structure. First, we investigate the filter PCP at the top to get an overview of the different value ranges. We notice the wide value range of the droplet quantities due to single outliers, which leads to the majority of droplets being squashed together for some of the quantities, e.g., this becomes quite apparent at the volume axis. We use the filter markers on the axis to remove the single droplet with very high volume, the jet base, and also omit droplets with an error flag for being too small—our expert specified that droplets below a threshold of ≈15 cells show unphysical behavior—or not being part of a trace. With the filtered data being rescaled on each axis within the SPLOM and second PCP, we can now see details and structures lost in the first overview PCP.

The first noticeable correlation we see is within the distance between droplet and jet center to velocity scatterplot (Fig. 5). In particular, we identified two main clusters, which are mostly symmetric to the diagonal: one cluster with high distance and high velocity (Fig. 5a) and a second cluster with small distance and small velocity (Fig. 6b). Furthermore, there are a few more outliers (Fig. 5c). With the first two main clusters, we conclude that the droplets in the center of the jet are slower than the droplets farther away from the center. We compare the scatterplots with the corresponding droplets in the 3D surface view to provide spatial context (Fig. 5d–f). In addition, the 3D droplets are colored by velocity using the Viridis colormap. The reason for these two clusters is that our dataset has a two-nozzle setup where we have an inner and outer zone with a different velocity at the boundary of the simulation domain. This can be easily seen by looking
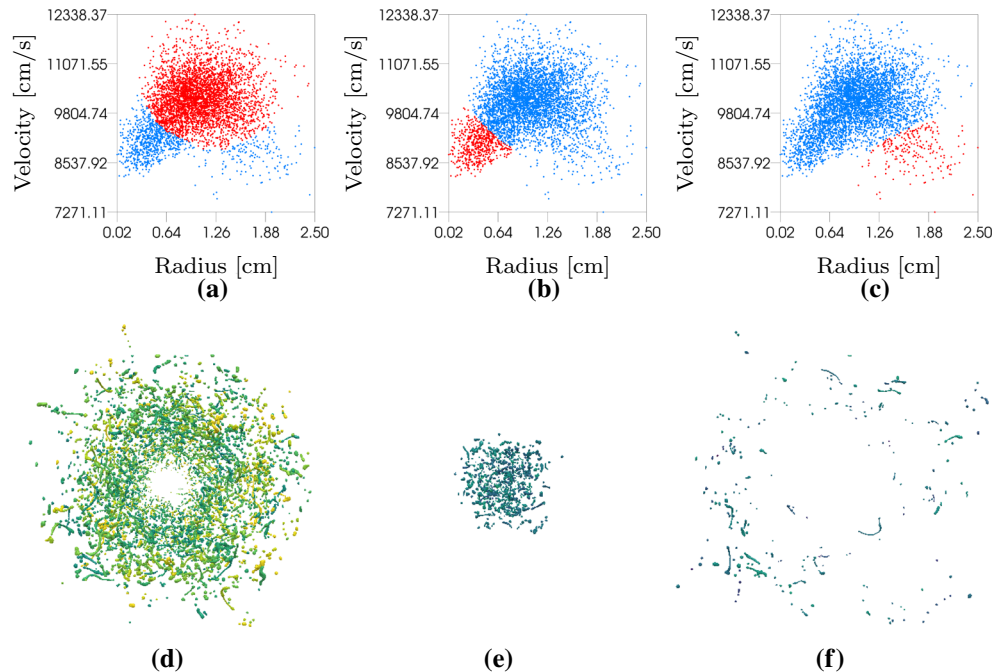

Fig. 5 Our analysis of the distance between droplet and jet center (here called radius) to velocity relationship. The top row shows the scatterplot with different selected data points (red). The linked 3D view below provides a spatial context for selection
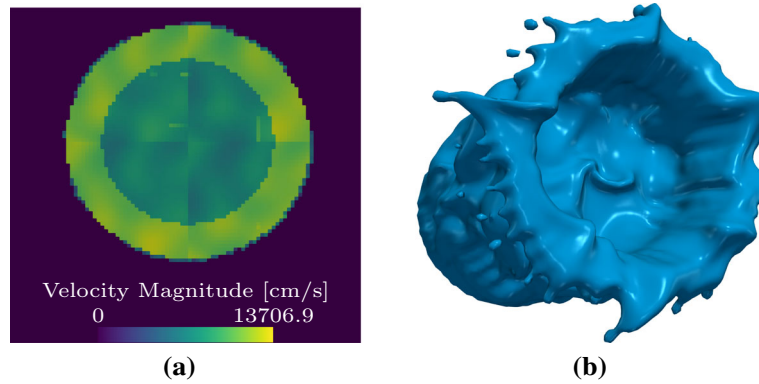
Fig. 6 a Inflow velocity magnitude profile at jet nozzle depicting the two-nozzle setup. **b** Isosurface of the jet base in the first simulation time step, reflecting the influence of the two-nozzle setup on the jet
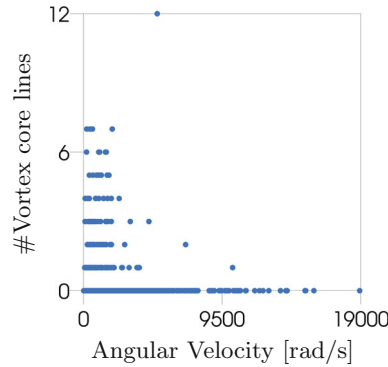

Fig. 7 A comparison of angular velocity to the number of vortex coreline segments within a droplet. These two quantities behave inversely, i.e., higher velocity yields a lower number of coreline segments
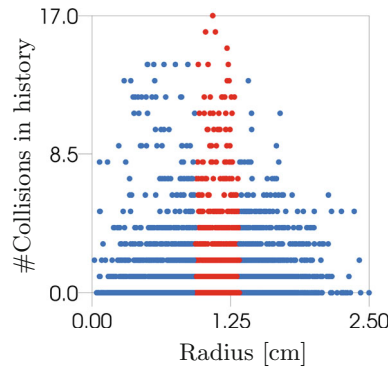

Fig. 8 A comparison of the radius of the droplet position around the jet center to the number of collision events in the history of a single droplet. Highlighted (red) is the radius range with a peak in the number of collision events

at the velocity magnitudes and the surface of the jet base in the first time step of the simulation (Fig. 6). However, it has previously been unclear how this setup affects droplet velocities further within the simulation domain.

Next, we look at the correlation between angular velocity and the number of vortex coreline segments (Fig. 7). Originally, our hypothesis was that a high angular velocity in the sense of rigid body rotation could be seen as a vortex, but in fact the plot shows exactly the opposite. We see that only droplets with a relatively small angular velocity have a high number of vortex core line segments, while droplets that have

at least one vortex core line segment have a relatively low angular velocity. We consider this to be an interesting finding, but were not yet able to confirm a physical explanation for this phenomenon.

The scatterplot of the distance between droplet and jet center compared to the number of merge events in the history of the jet also looks interesting (Fig. 8). It exhibits a triangular shape, except for a few outliers. The most inner and outer droplets seem to have only few merge events, while there is a ring of droplets with average distance, that appears to have a very high number of such occurrences. We attribute this to the dual nozzle injection, where slower droplets in the inner and faster droplets in the outer ring collide in a transition area. The peak in the number of merge events accordingly indicates this contact area.

5.2 Clustering

The clustering results depict different characteristics (cf. Fig. 9). We notice that the volume is one of the main factors, which separates these clusters. It further shows in Fig. 9b that the velocity is a factor orthogonal to the volume. In Fig. 9c, we can see the contour of a cluster with respect to the spherical anisotropy. This is especially remarkable because this value was not used for the clustering. Accordingly, we assume that this is an indicator for the correlation of quantities. A limitation may be that there is no direct physical interpretation of these clusters. While it would be a great result if we had found one, our clustering method is based on the trace average of the extracted quantities. This is a very simplified projection of a droplet, probably not covering all physical laws and may not be entirely based on intrinsic physical quantities. Therefore, the clusters could only have a phenomenological interpretation by looking at the quantity distribution within the SPLOM view as sketched by Fig. 9. Nevertheless, we think these are still structural and dynamical relevant clusters representing groups of similar droplet behavior within the abstract quantity space.

Looking at the 3D surface view, we find that each cluster has similar droplets in reference to its shape and size (Fig. 10). Overall this clustering gives us an overview of the different types of droplets within this dataset. It provides additional information to statistical quantity value distribution, for example, within the PCP view, by addressing higher-order connection in the high-dimensional space of the quantity values. However, remember that we neglect temporal effects as we averaged traces before clustering (cf. Sect. 3.3).

5.3 Droplet anomaly-guided analysis

Finally, we analyze the results based on prediction anomaly. In Fig. 11a, we see a correlation between the angular velocity and the momentum, with most data points being close to the axis. That means a droplet generally exhibits high velocity or high momentum, but not both. By coloring the data points by the trace prediction error, we see that the further droplets are away from this trend along the axis, the higher is the prediction error. In the bottom left, we see a dense region with mostly low anomaly droplets. Zooming reveals a few droplets with a high anomaly, which attract our attention (Fig. 11b). Using brushing on the highest anomaly data point in this region (marked red) and observing it in the linked 3D surface and flow view, we see that this droplet is located at the boundary of the domain. Identifying such problematic cases is important for studying edge cases in the simulation and discarding them from further consideration.
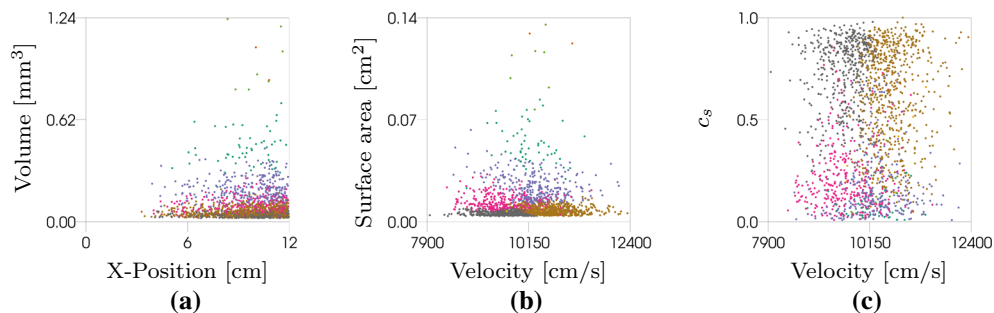

Fig. 9 Data points are colored by cluster ID from the eight-class hierarchical ward clustering. **a** Position in x-direction (flow direction of the jet) compared to volume. Notice the relation of the clusters for higher volumes. For lower volumes, clusters are not distinguishable. **b** Comparing velocity to surface, we can see separation of clusters in both dimensions. **c** Same holds for comparing velocity to spherical anisotropy (anisotropy is not considered during clustering)
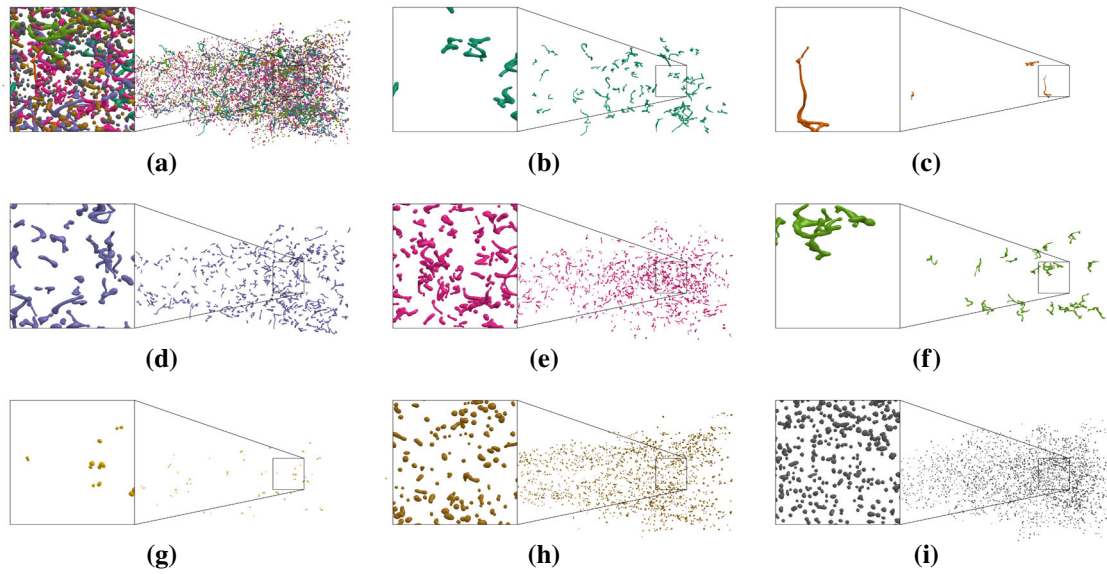
Fig. 10 Same clusters as in Fig. 10, but within the 3D surface view. **a** All clusters in one view. **b–i** Each cluster on its own. Note the similarity of droplet surface shape and size within clusters
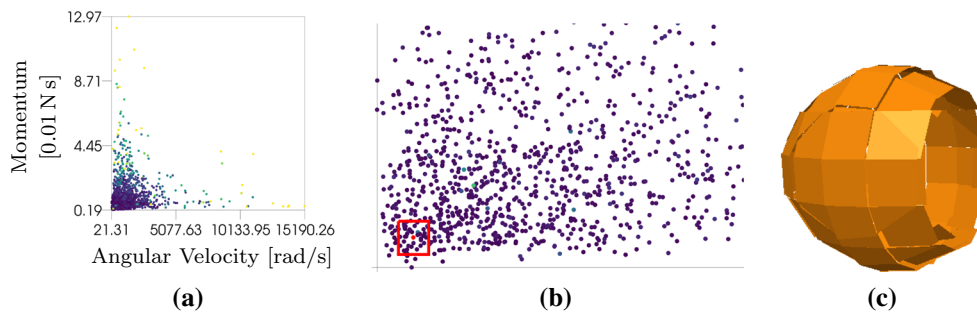

Fig. 11 **a** The comparison of angular velocity to momentum shows a trend of data points aligning along the axis. Coloring by prediction anomaly seems to correlate with outliers of this trend. **b** Zoom to bottom left of **a**. Selected (red) is the data point with the highest anomaly in this range. **c** 3D surface view of the selected droplet in **b**. We can see the domain boundary cuts this droplet. This implies to be the source for this droplet being an anomaly

Figure 4a shows the 3D surface view depicting time step 410. Color depicts our anomaly estimation $\bar{\varepsilon}$, with higher anomaly indicated by reddish colors and lower anomaly by yellowish ones. Gray indicates droplet instances that are within the $w - 1$ first time steps of a trace and thus do not provide prediction nor anomaly information $\bar{\varepsilon}$ (this always applies to the jet itself, it is always at the beginning of the main trace due to droplets splitting from it continuously). To reduce occlusion and clutter, a typical first step is to omit those structures, which can be accomplished by requiring $\bar{\varepsilon} \geq 0$, because $\bar{\varepsilon}$ is set to a negative constant in our implementation if no $\bar{\varepsilon}$ value can be computed (cf. Fig. 12a). Here, we observe that ligaments, i.e., the elongated droplets that typically break up into smaller droplets, later on, exhibit very high $\bar{\varepsilon}$, i.e., they show temporal behavior that strongly deviates from the trends of the majority of the droplets. We also note that the structures with the highest $\bar{\varepsilon}$ are typically tiny droplets (Fig. 13), which led us to the hypothesis that the quantities that we computed suffered from discretization artifacts for small droplets, even if they are still larger than the required minimum size initially indicated by our domain scientists. In particular, such droplets tend to exhibit alternating $\bar{\varepsilon}$, switching between high and low values at a high temporal frequency, which supports our hypothesis of discretization issues due to insufficient resolution. We also observed such alternating behavior of $\bar{\varepsilon}$ for very thin ligaments. As a result, we do not consider very small droplets nor ligaments with this analysis component in the following. To accomplish this, we filter by spherical anisotropy c_s, requiring $c_s > 0.4$ in this case. Additionally, we suppress all droplets with low anomaly by adjusting our filter to $\bar{\varepsilon} \geq 0.15$. To obtain a better space-time overview, we now enable the simultaneous
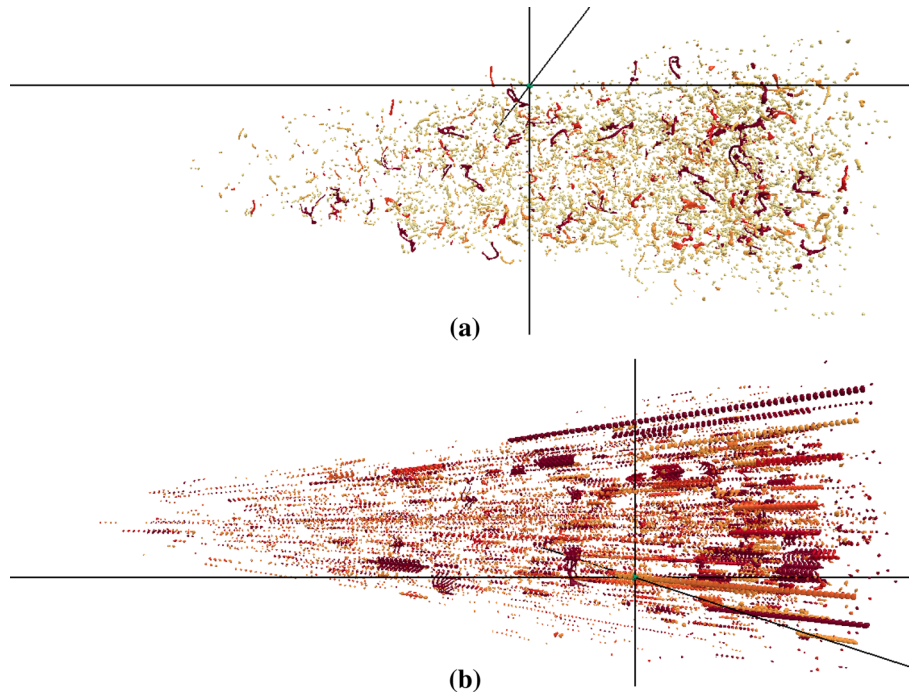
(a)

(b)

Fig. 12 Exploration of the dataset within the 3D surface view (cf. Figure 4a (*iv*)). After loading the dataset, the time filter is set to show only a single time step and a second filter omits all droplet instances that lack total error (our anomaly indicator) $\bar{\varepsilon}$, leading to the view in **a**. The user now selects a droplet of interest for further analysis (black cross). Here, the droplet shown in Figure 14 (left) is selected. **b** Further example. Now the time filter is set to aggregate all time steps of the dataset, leading to dotted line-like structures showing the same droplet over time. To further reduce visual clutter, filters are used to show only droplets with anomaly measure $\bar{\varepsilon} > 0.15$ and for suppressing ligaments by requiring spherical anisotropy $c_s > 0.4$. Here, the droplet in Fig. 15 is selected
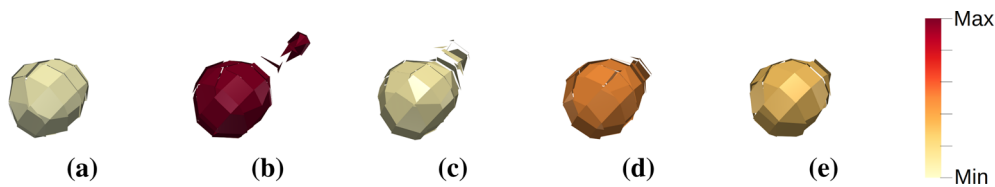

(a)　　　**(b)**　　　**(c)**　　　**(d)**　　　**(e)**

Fig. 13 Droplet exhibiting the highest anomaly. Structures with the highest anomaly measure $\bar{\varepsilon}$ are tiny droplets whose quantity computation suffers from discretization issues. In **b** the collision with a smaller droplet artifact can be observed, which was missed by the trace generation (cf. Sect. 3.2). Further, due to the nature of PLIC, the strongly discontinuous surface reconstruction can be seen in the form of gaps (cf. Sect. 3.1). Successive time steps **a**–**e** exhibit alternating behavior of $\bar{\varepsilon}$, supporting the hypothesis of insufficient resolution of the simulation grid

display of all time steps with the chosen filtering criteria, leading to Fig. 12b. This way, we observe many long traces in the 3D surface view, which provides the point of origin for our further investigations.

We now pick a trace with exceptionally high anomaly estimation (selection indicated by a black cross in Fig. 12a). A closer inspection of that trace in the time-view (Fig. 4a (*v*)) reveals that the droplet becomes rounder over time and that the anomaly indicator $\bar{\varepsilon}$ stays almost constant over time. This observation turned out to be quite rare, since typically, the anomaly measure reduces quite quickly, particularly if the respective droplet becomes roundish. A thorough investigation of the respective plots of the original traces and the individual deviations of the predicted quantities from the original ones did not provide insights on the causes of this behavior. We started to hypothesize that the internal flow within the droplet might provide insights. We thus initiated flow visualization of the liquid phase of the droplet (Fig. 14) (left). Note that we use a linear mapping of velocity magnitude to glyph color, but a logarithmic mapping of velocity magnitude to glyph size within Fig. 14 and following. Interestingly, we observed a distinguished and strong saddle-type flow pattern in the internal flow, in the frame of reference moving with the velocity of the center of mass of the droplet. We investigated other cases with high anomaly measure $\bar{\varepsilon}$, either by direct selection in the 3D

High Anomaly Droplet case.　　　High Anomaly Droplet II case.
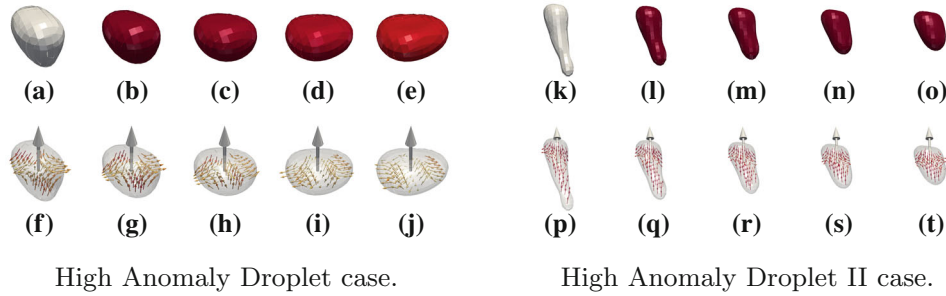
Fig. 14 Left: Selected droplet from Fig. 12a. We found that droplets with low temporal decay of $\bar{\varepsilon}$ exhibit distinguished saddle-type flow patterns in their interior flow, in the frame of reference moving with the center of mass of the droplet. **Right:** This is the droplet found most similar by similarity search in Fig. 4a (*vi*)

view or using our similarity search, as provided in the lower half of our tool. Interestingly, most cases of non-ligament (more or less roundish) droplets with high $\bar{\varepsilon}$ turned out to exhibit such saddle-type flow patterns in their interior, including the case shown in Fig. 14 (right), found as most similar to that droplet by similarity search.

This motivated us to investigate droplets with different anomaly behavior. While decaying $\bar{\varepsilon}$ is quite common, we wanted to look at droplets for which $\bar{\varepsilon}$ increases over time. The hypothesis behind this reasoning is that a collision of droplets often causes anomalous droplet behavior, but while the products of the collision move over time, they tend to "calm down" and become better predictable by our regression approach, and thus $\bar{\varepsilon}$ decays. In Fig. 12b, we were able to identify such a droplet, select it by picking, and investigate its temporal behavior in the time view. Detailed flow analysis is shown in Fig. 15. Interestingly, this droplet has been the only one we could find with a strong vortex in its interior.

Finally, for comparison and validation of the utility of our anomaly estimation, we investigated droplets with low anomaly measure $\bar{\varepsilon}$. We selected them either by direct picking or by similarity search. The majority of droplets with low $\bar{\varepsilon}$ exhibit shear flow in its interior, as shown in Fig. 16 (left). We were able to find only one case with deviating flow pattern. This droplet (Fig. 16) (right) exhibits a "half" saddle-type flow pattern (imagine splitting a 3D saddle along its 2D manifold) but with low flow velocity in the frame of reference moving with the observer.

6 Discussion and future work

We developed a visual analysis approach for investigating droplet characteristics and behaviors occurring in large ensembles such as sprays. It allows exploring what initially are terabytes of simulation data depicting complex small-scale effects with a multi-stage workflow. On the one hand, we reduced the data by
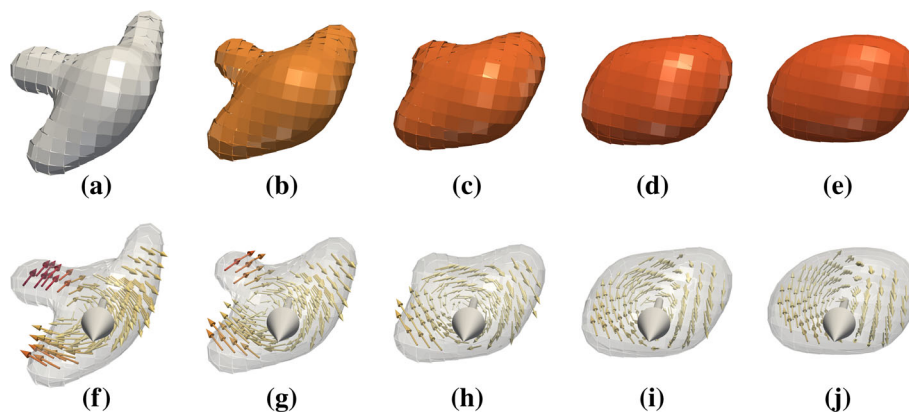

Fig. 15 An increasing anomaly droplet case. The scarce case of temporally increasing $\bar{\varepsilon}$ brought our attention to this case, which, in the frame of reference moving with the center of mass of the droplet, exhibits a strong vortex in its interior flow. This droplet has been the only one we could identify to contain a vortex
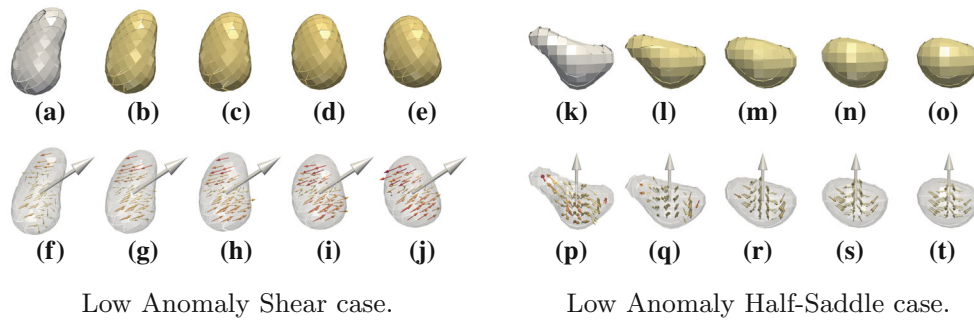
Low Anomaly Shear case.   Low Anomaly Half-Saddle case.

Fig. 16 **Left:** Droplets with low anomaly measure $\bar{\varepsilon}$ generally exhibit shear flow patterns in their interior, in frame of reference moving with the center of mass of the droplet. **Right:** The only exception to our hypothesis, that low anomaly droplets exhibit shear flow patterns in the interior. This droplet exhibits a weak "half-saddle" pattern, i.e., half of a 3D saddle, similar to a detachment point in its interior, in the frame of reference moving with the droplet

extracting droplet surfaces and physical quantities. This allowed us to give an overview and investigate large numbers of droplets at once during interactive exploration. The physical quantities were further used in advanced analysis steps to (1) get an overview of different kinds of droplets via clustering (where we introduced a domain-specific optimization to be able to handle a large number of droplets), and (2) identify droplets with irregular temporal behavior by adapting a recent machine learning-based method. On the other hand, the complete original data were still available for the detailed investigation of selected cases. We integrated different analysis components to be able to provide the user with the full breadth of required interactive analysis functionality from investigating relationships of extracted physical quantities, droplet traces, and results from clustering and ML-based analysis to a spatial overview, and finally, a direct link to ParaView supporting advanced flow analysis for the detailed investigation of selected droplets.

So far, we have applied our tool to gain insights from a single large-scale direct numerical simulation datasets of two-phase flow of non-Newtonian jets. However, we believe our approach directly generalizes to various simulations of this kind. We have chosen basic physical quantities based on raw simulation data from the VOF method (and accordingly will be available from nearly all VOF-based simulations). Depending on the simulation results, the list of properties could be extended, e.g., if pressure is available, or a higher resolution allows for a precise calculation of derivation-based quantities. To test the stability of our quantity calculation, we have downsampled the simulation grid of each droplet by merging a cube of $2 \times 2 \times 2$ cells into a single cell and repeated the quantity calculation thereon. As expected, the calculations are quite stable for large droplets containing many grid cells and less stable for smaller droplets consisting of fewer cells. An example for this behavior in the form of the velocity value is shown in Fig. 17a. Further, we can look at the temporal development of droplet quantities. Due to the relatively high temporal resolution of our dataset, we expect smooth changes in the quantities. Figure 17b shows that this only holds for the high-resolution grid, while the downsampled grid introduces numerical errors due to the coarse grid.

We would generally expect the ANN-based anomaly detection to work likewise for other data. However, as our prediction is based on traces, we need a minimum temporal resolution, especially some highly chaotic
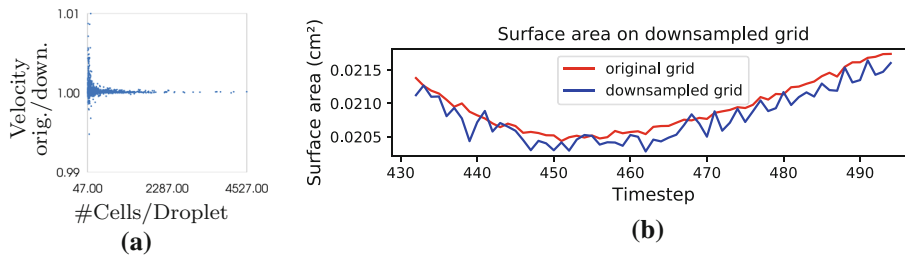

Fig. 17 An analysis of the quantity calculation on the original grid vs. downsampled grid. **a** Quotient of velocity on the original and downsampled grid, plotted over the number of cells per droplet. Note that for a large number of cells per droplet, this quotient is near one, indicating the quantity calculation is stable over different grid resolutions, while for a low number of cells, we see diverging calculation results. **b** Surface area over time for the droplet in Fig. 14. The number of cells for this droplet on the original grid varies between 811 and 862 cells. The surface area changes are smooth on the original grid but less stable on the downsampled grid

datasets might be problematic, where splits and merges happen in nearly every time step for each droplet. That being said, we generally consider jets to be already among the more difficult real-world examples. However, the trace aggregation approach we use for hierarchical clustering might not adequately transfer to other examples. While the representation of traces through aggregates proved useful in our analysis scenario, more elaborate, adaptive approaches are potentially required in other contexts.

In future work, we aim to study additional variants of jets, and beyond this, we will explore possibilities to generalize our technique for types of simulation data in which large amounts of small entities need to be investigated. We plan to incorporate further split and merge events to understand the atomization process better. While we already collect the number of these events in the entire history of a droplet, we also aim to integrate a graph view in our tool (enriched with our extracted quantities) and extend our ML-based anomaly detection to consider split events and merge events explicitly. Finally, the main goal would be to allow a user to directly influence our clustering and machine learning component during interactive analysis for specified subsets or partitions of the data.

Appendix A

A.1 Definition of Physically Motivated Quantities

Droplet volume: We compute the volume of each droplet by integrating the f-field. That is, the volume $\overline{\mu}_\iota$ of the droplet with identifier ι is obtained by integrating the product of the f-field and the cell volume for all cells i where $l_i = \iota$:

$$\overline{\mu}_\iota := \sum_{i \in \mathcal{C}_\iota} f_i \mu_i \,, \tag{3}$$

with $\mathcal{C}_\iota := \{j \mid l_j = \iota\}$, and μ_i being the volume of cell i.

Droplet area-to-volume ratio: As motivated above, we compute the area $\overline{A}_\iota$ of droplet ι by integrating the area of the respective PLIC patches, i.e.,

$$\overline{A}_\iota := \sum_{i \in \mathcal{C}_\iota} A_i \,, \tag{4}$$

with A_i being the area of the PLIC patch in cell i, i.e., $A_i = 0$ if $f_i = 0$ or $f_i = 1$. From that, we compute the area-to-volume ratio $\overline{\alpha}_\iota$ (Karch et al. 2018) of the droplet as

$$\overline{\alpha}_\iota := \frac{\overline{A}_\iota}{\overline{\mu}_\iota} \frac{r_s}{3} \,, \tag{5}$$

with $r_s := \sqrt[3]{(3\overline{\mu}_\iota)/(4\pi)}$ being the radius of a sphere with volume $\overline{\mu}_\iota$.

Droplet velocity and momentum: The velocity of the center of mass of droplet ι follows to be

$$\overline{\boldsymbol{u}}_\iota := \frac{1}{\overline{\mu}_\iota} \sum_{i \in \mathcal{C}_\iota} m_i \boldsymbol{u}_i \,, \tag{6}$$

with cell-based flow velocity $\mathbf{u}_i$ in cell i, and $m_i := f_i \mu_i$. From that, the droplet's momentum computes

$$\overline{\boldsymbol{p}}_\iota := \overline{\mu}_\iota \overline{\boldsymbol{u}}_\iota \,. \tag{7}$$

Notice that we assume the density of the liquid to be one because the density is often not provided explicitly for the liquid phase in two-phase simulations (as in our case), and since density represents only a scaling factor (note that liquids are generally treated as being incompressible). As a consequence, the total mass of droplet ι equals its volume $\overline{\mu}_\iota$.

Auxiliary measures: A common measure in astrophysics and particle systems is the center of mass. It provides a frame of reference and enables the computation of derived quantities. The center of mass $\overline{\mathbf{x}}_\iota$ of droplet ι computes

$$\overline{\boldsymbol{x}}_\iota := \frac{1}{\overline{\mu}_\iota} \sum_{i \in \mathcal{C}_\iota} m_i \boldsymbol{x}_i \,, \tag{8}$$

with $\mathbf{x}_i$ being the center of cell i.

From that, it is a common step to compute the total angular momentum for a set of particles, resulting in our case in the angular momentum $\overline{\mathbf{L}}_\iota$ of the droplet:

$$\overline{\boldsymbol{L}}_\iota := \sum_{i \in \mathcal{C}_\iota} \hat{\boldsymbol{x}}_i \times m_i \boldsymbol{u}_i \,, \tag{9}$$

with $\hat{\mathbf{x}}_i := \mathbf{x}_i - \overline{\mathbf{x}}_\iota$ being the center of cell i relative to $\overline{\mathbf{x}}_\iota$. Notice that, for the example of particle systems, this total angular momentum describes the rotational motion of the "rigid-body aspect" of the particle set, or in other words, the rotational motion of the entirety of the particles. Thus, it represents the "rigid-body rotation component" of the droplet in terms of the flow of the liquid in our context.

In analogy, the droplet's inertia tensor $\overline{\Theta}_\iota$ computes

$$\overline{\Theta}_\iota := \begin{pmatrix} \sum_i m_i(\hat{y}_i^2 + \hat{z}_i^2) & -\sum_i m_i\hat{y}_i\hat{x}_i & -\sum_i m_i\hat{z}_i\hat{x}_i \\ -\sum_i m_i\hat{x}_i\hat{y}_i & \sum_i m_i(\hat{x}_i^2 + \hat{z}_i^2) & -\sum_i m_i\hat{z}_i\hat{y}_i \\ -\sum_i m_i\hat{x}_i\hat{z}_i & -\sum_i m_i\hat{y}_i\hat{z}_i & \sum_i m_i(\hat{x}_i^2 + \hat{y}_i^2) \end{pmatrix} \,, \tag{10}$$

with $\hat{\mathbf{x}}_i =: (\hat{x}_i, \hat{y}_i, \hat{z}_i)^\top$ and $i \in \mathcal{C}_\iota$. From that, the angular velocity $\overline{\omega}_\iota$ of droplet ι is obtained by

$$\overline{\omega}_\iota := \overline{\Theta}_\iota^{-1} \overline{\mathbf{L}}_\iota \,. \tag{11}$$

Notice that $\overline{\Theta}_\iota$ can become singular for very small droplets, which is one reason we exclude such very small droplets from our analysis.

Total droplet energy: The total energy E_ι of droplet ι (notice that the addressed simulations exclude thermodynamical and chemical processes) computes

$$E_\iota := \sum_{i \in \mathcal{C}_\iota} \frac{1}{2} m_i \boldsymbol{u}_i^2 \,. \tag{12}$$

This total energy can be decomposed into kinetic $E_\iota^{\mathbf{u}}$, rotational E_ι^ω, and residual energy E_ι^δ:

$$E_\iota := E_\iota^{\mathbf{u}} + E_\iota^\omega + E_\iota^\delta \,, \tag{13}$$

as explained next.

Kinetic droplet energy: The kinetic energy $E_\iota^{\mathbf{u}}$ of droplet ι with respect to its velocity $\overline{\mathbf{u}}_\iota$ is

$$E_\iota^{\mathbf{u}} := \frac{1}{2} \overline{\mu}_\iota \overline{\mathbf{u}}_\iota^2 \,. \tag{14}$$

This measure represents the kinetic energy of the rigid-body property of the liquid component representing a droplet.

Rotational droplet energy: The rotational energy E_ι^ω of droplet ι computes

$$E_\iota^\omega := \frac{1}{2}\overline{\varpi}_\iota^\top \overline{\Theta}_\iota \overline{\varpi}_\iota \,. \tag{15}$$

This rotational energy captures the rigid-body part of the liquid dynamics of the droplet with respect to rotational motion.

Residual droplet energy: The residual energy E_ι^δ of droplet ι finally computes

$$E_\iota^\delta := E_\iota - E_\iota^{\mathbf{u}} - E_\iota^\omega \,, \tag{16}$$

and includes deformation of the droplet, as well as the non-rigid flow in its interior. In that sense, E_ι^δ captures the deviations of a droplet's flow from rigid-body dynamics. Due to limited numerical accuracy, E_ι^δ can turn out to be slightly negative, in which case we clamp it to zero and set $E_\iota^\omega := E_\iota - E_\iota^{\mathbf{u}}$.

References

Aigner W, Bertone A, Miksch S, Tominski C, Schumann H (2007) Towards a conceptual framework for visual analytics of time and time-oriented data. Winter Simul Conf. https://doi.org/10.1109/WSC.2007.4419666

Ayachit U (2015) The para view guide: a parallel visualization application. Kitware Inc, New York

Bai Z, Brunton SL, Brunton BW, Kutz JN, Kaiser E, Spohn A, Noack BR (2017) Data-driven methods in fluid dynamics: sparse classification from experimental data. Springer International Publishing, Cham, pp 323–342. https://doi.org/10.1007/978-3-319-41217-7_17

Basser PJ, Pierpaoli C (2011) Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor mri. J Magn Reson 213(2):560–570. https://doi.org/10.1016/j.jmr.2011.09.022

Bengio Y, Courville A, Vincent P (2013) Representation learning: a review and new perspectives. IEEE Trans Pattern Anal Mach Intell 35(8):1798–1828. https://doi.org/10.1109/TPAMI.2013.50

Biswas A, Ahrens J, Dutta S, Musser J, Almgren A, Turton TL (2020) Feature analysis, tracking, and data reduction: an application to multiphase reactor simulation mfix-exa for in-situ use case. Comput Sci Eng. https://doi.org/10.1109/MCSE.2020.3016927

Bürger R, Muigg P, Ilcík M, Doleisch H, Hauser H (2007) Integrating local feature detectors in the interactive visual analysis of flow simulation data. In: Museth K, Moeller T, Ynnerman A (eds) Eurographics/ IEEE-VGTC symposium on visualization, the Eurographics Association. Springer, Berlin. https://doi.org/10.2312/VisSym/EuroVis07/171-178

Doleisch H, Gasser M, Hauser H (2003) Interactive feature specification for focus+context visualization of complex simulation data. In: Bonneau GP, Hahmann S, Hansen CD (eds) Eurographics / IEEE VGTC symposium on visualization, the Eurographics Association. Springer, Berlin. https://doi.org/10.2312/VisSym/VisSym03/239-248

Doleisch H, Mayer M, Gasser M, Wanker R, Hauser H (2004a) Case study: visual analysis of complex, time-dependent simulation results of a diesel exhaust system. In: Deussen O, Hansen C, Keim D, Saupe D (eds) Eurographics / IEEE VGTC symposium on visualization, the Eurographics Association. Springer, Berlin. https://doi.org/10.2312/VisSym/VisSym04/091-096

Doleisch H, Muigg P, Hauser H (2004b) Interactive visual analysis of hurricane Isabel with SimVis (2004b). http://vis.computer.org/vis2004contest/vrvis/

Dutta S, Shen H (2016) Distribution driven extraction and tracking of features for time-varying data analysis. IEEE Trans Vis Comput Graph 22(1):837–846. https://doi.org/10.1109/TVCG.2015.2467436

Eisenschmidt K, Ertl M, Gomaa H, Kieffer-Roth C, Meister C, Rauschenberger P, Reitzle M, Schlottke K, Weigand B (2016) Direct numerical simulations for multiphase flows: an overview of the multiphase code fs3d. Appl Math Comput 272:508–517. https://doi.org/10.1016/j.amc.2015.05.095

Endert A, Ribarsky W, Turkay C, Wong BLW, Nabney I, Blanco ID, Rossi F (2017) The state of the art in integrating machine learning into visual analytics. Comput Graph Forum 36(8):458–486. https://doi.org/10.1111/cgf.13092

Ertl M (2019) Direct numerical investigations of non-newtonian drop oscillations and jet breakup. PhD thesis, University of Stuttgart, https://doi.org/10.18419/opus-10764

Ertl M, Weigand B (2017) Analysis methods for direct numerical simulations of primary breakup of shear-thinning liquid jets. At Sprays 27(4):303–317. https://doi.org/10.1615/AtomizSpr.2017017448

Ester M, Kriegel HP, Sander J, Xu X (1996) A density-based algorithm for discovering clusters in large spatial databases with noise. In: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, AAAI Press, Portland, Oregon, KDD'96, pp 226–231

Garth C, Tricoche X, Scheuermann G (2004) Tracking of vector field singularities in unstructured 3d time-dependent datasets. IEEE Vis 2004:329–336. https://doi.org/10.1109/VISUAL.2004.107

Glorot X, Bordes A, Bengio Y (2011) Deep sparse rectifier neural networks. In: Gordon G, Dunson D, Dudík M (eds) Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, JMLR Workshop and Conference Proceedings, Fort Lauderdale, FL, USA, Proceedings of Machine Learning Research, pp 315–323

Gralka P, Becher M, Braun M, Frieß F, Müller C, Rau T, Schatz K, Schulz C, Krone M, Reina G, Ertl T (2019) MegaMol: a comprehensive prototyping framework for visualizations. Eur Phys J Spec Top 227(14):1817–1829. https://doi.org/10.1140/epjst/e2019-800167-5

Heinemann M (2018) Ml-based visual analysis of droplet behaviour in multiphase flow simulations. Master's thesis, University of Stuttgart, https://doi.org/10.18419/opus-10109

Hirt CW, Nichols BD (1981) Volume of fluid (vof) method for the dynamics of free boundaries. J Comput Phys 39(1):201–225. https://doi.org/10.1016/0021-9991(81)90145-5

Johnson SC (1967) Hierarchical clustering schemes. Psychometrika 32(3):241–254. https://doi.org/10.1007/BF02289588

Karch GK, Sadlo F, Meister C, Rauschenberger P, Eisenschmidt K, Weigand B, Ertl T (2013) Visualization of piecewise linear interface calculation. IEEE Pac Vis Symp (PacificVis). https://doi.org/10.1109/PacificVis.2013.6596136

Karch GK, Beck F, Ertl M, Meister C, Schulte K, Weigand B, Ertl T, Sadlo F (2018) Visual analysis of inclusion dynamics in two-phase flow. IEEE Trans Vis Comput Graph 24(5):1841–1855. https://doi.org/10.1109/TVCG.2017.2692781

Kingma DP, Ba J (2014) Adam: A method for stochastic optimization. In: 3rd International Conference on Learning Representations, ICLR, Conference Track Proceedings

Post FH, Vrolijk B, Hauser H, Laramee RS, Doleisch H (2003) The state of the art in flow visualisation: feature extraction and tracking. Comput Graph Forum 22(4):775–792. https://doi.org/10.1111/j.1467-8659.2003.00723.x

Rosenberger G, Nestor PG, Oh JS, Levitt JJ, Kindleman G, Bouix S, Fitzsimmons J, Niznikiewicz M, Westin CF, Kikinis R, McCarley RW, Shenton ME, Kubicki M (2012) Anterior limb of the internal capsule in schizophrenia: a diffusion tensor tractography study. Brain Imaging Behav 6(3):417–425. https://doi.org/10.1007/s11682-012-9152-9

Shi K, Theisel H, Hauser H, Weinkauf T, Matkovic K, Hege HC, Seidel HP (2009) Path line attributes: an information visualization approach to analyzing the dynamic behavior of 3D time-dependent flow fields. Springer, Berlin, pp 75–88. https://doi.org/10.1007/978-3-540-88606-8_6

Straub A, Ertl T (2020) Visualization techniques for droplet interfaces and multiphase flow. In: Lamanna G, Tonini S, Cossali GE, Weigand B (eds) Droplet interactions and spray processes. Springer International Publishing, Cham, pp 203–214. https://doi.org/10.1007/978-3-030-33338-6_16

Straub A, Heinemann M, Ertl T (2019) Visualization and visual analysis for multiphase flow. In: Cossali G, Tonini S (eds) Proceedings of the DIPSI workshop 2019: droplet impact phenomena and spray investigations. Università degli studi di Bergamo, Bergamo, pp 75–88. https://doi.org/10.6092/DIPSI2019_pp25-27

Sujudi D, Haimes R (1995) Identification of swirling flow in 3-d vector fields. Comput Fluid Dyn Conf. https://doi.org/10.2514/6.1995-1715

Theisel H, Seidel HP (2003) Feature flow fields. In: Bonneau GP, Hahmann S, Hansen CD (eds) Eurographics / IEEE VGTC symposium on visualization, the Eurographics Association. Springer, Berlin. https://doi.org/10.2312/VisSym/VisSym03/141-148

Tkachev G, Frey S, Ertl T (2019) Local prediction models for spatiotemporal volume visualization. IEEE Trans Vis Comput Graph. https://doi.org/10.1109/TVCG.2019.2961893

Tompson J, Schlachter K, Sprechmann P, Perlin K (2017) Accelerating Eulerian fluid simulation with convolutional networks. In: Precup D, Teh YW (eds) Proceedings of the 34th International Conference on Machine Learning, PMLR, International Convention Centre, Sydney, Australia, Proceedings of Machine Learning Research, pp 3424–3433

Tzeng Fan-Yin, Ma Kwan-Liu (2005) Intelligent feature extraction and tracking for visualizing large-scale 4d flow simulations. Proc ACM/IEEE Conf Super Comput. https://doi.org/10.1109/SC.2005.37

Youngs DL (1984) An interface tracking method for a 3D Eulerian hydrodynamics code. Atomic Weapons Research Establishment (AWRE). Tech Rep 44(92):35