

# Quantitative characterization of hydrophobic agglomeration at different mixing intensities using a copula-based probabilistic modeling approach

Niklas Eiermann<sup>1,\*</sup>, Orkun Furat<sup>1,\*</sup>, Jan Nicklas<sup>2</sup>, Urs A. Peuker<sup>2</sup>, Volker Schmidt<sup>1</sup>

<sup>1</sup>*Institute of Stochastics, Ulm University, Helmholtzstraße 18, 89069 Ulm, Germany*

<sup>2</sup>*Institute of Mechanical Process Engineering and Mineral Processing, Technische Universität Bergakademie Freiberg, Agricolastraße 1, 09599 Freiberg, Germany*

## Abstract

The agglomeration of small poorly wetted alumina particles in a stirred tank is investigated. For different experimental conditions, two bivariate probability densities for the area-equivalent diameter and aspect ratio of primary particles and agglomerates, respectively, are determined, using 2D image data from an inline camera system. Throughout each experiment, these densities do not change since the geometries of primary particles are unaffected by the experimental conditions, while large agglomerates fragment into multiple smaller ones, which results in an equilibrium state regarding the distribution of agglomerate descriptors. Mixtures of these densities are used to model the contents of the stirred tank at each time step of the experiments. Analytical functions, whose parameters characterize the agglomeration dynamics, are fitted to the time-dependent weights of these mixtures. This enables a quantitative comparison of agglomeration processes, highlighting the impact of mixing intensity on the joint distribution of agglomerate descriptors.

**Keywords:** Hydrophobic agglomeration, Inline probe, Area-equivalent diameter, Aspect ratio, Bivariate probability density, Archimedean copula.

## 1 Introduction

The application case comes from liquid metal filtration, since undesired inclusion particles agglomerate through turbulent collision, forming larger, multi-component particles that are easier to remove [1]. The increased size of agglomerated inclusion particles leads to augmented particle deposition at the inner surface of ceramic depth filters, which are utilized for the cleaning of molten metal from inclusion particles [2]. This enables metallurgists to optimize steelmaking for cleaner and more durable high-quality products. Since metal melts are opaque and high temperatures are necessary for maintaining their molten state, it is impractical to study agglomeration processes in melts. Therefore, aqueous sodium chloride (NaCl) solutions in combination with poorly wetted, i.e., hydrophobic, particles often serve as a model [3]. In the present paper, we investigate the agglomeration of silanized alumina particles within a stirring tank filled with 0.75 mol/l NaCl at the isoelectric point of pH 7.3. The silanization of the solid surface configures a wetting angle of  $134^\circ$ , which results in a poor wettability of the alumina particles by the electrolyte solution, similar to the poor wettability of non-metallic inclusion particles by molten metal [4]. Hydrophobic agglomerates are generally larger and more stable than hydrophilic ones, as nanobubbles on the surface of hydrophobic particles form capillary bridges [2, 4, 5]. Besides the purification of metals, agglomeration processes hold significance in the efficient separation of suspensions into liquid and particles [6]. Comprehending agglomeration dynamics involves studying factors such as the turbulent flow field, particle size, wettability, viscosity, and temperature.

The aim of the present paper is to complement the results which have been achieved in [4] for various agglomeration experiments, where mixtures of normal distributions have been fitted to single particle descriptors, e.g. the aspect ratio, of primary particles and agglomerates segmented from 2D image data measured in a stirred tank. More precisely, the status of agglomeration processes has been investigated at 52 time steps over the course of 47 minutes long batch agglomeration experiments. Specifically, the parameters of the underlying normal distributions of the

---

\* NE and OF contributed equally.

Email addresses: niklas.eiermann@uni-ulm.de (Niklas Eiermann), orkun.furat@uni-ulm.de (Orkun Furat)

particle descriptors of primary particles and agglomerates, respectively, as well as the mixing weights, i.e., the time-dependent number/volume-based fractions of agglomerates (or primary particles) of all particles in the stirred tank, have been estimated. One of the key observations made in [4] is that the parameters of the normal distributions remain rather constant throughout the 52 time steps, whereas the mixing weights significantly change over time. This indicates that the descriptors of primary particles as well as agglomerates follow fixed distributions, which do not depend on the current stage of the agglomeration process, whereas the ratio of primary particles to agglomerates changes along the process.

Building on the results obtained in [4], the present paper extends the probabilistic modeling approach of hydrophobic agglomeration based on bivariate data of two-dimensional particle descriptor vectors, simultaneously considering the area-equivalent diameter and the aspect ratio of primary particles and agglomerates. In particular, instead of focusing on one-dimensional normal distributions, a large variety of parametric probability distributions [7] is considered as candidates for the (marginal) distributions of single particle descriptors. Furthermore, the fact is taken into account that size and shape descriptors of particles can be correlated. For example, in some cases, small particles can be nearly spherical, whereas large particles are elongated. Thus, in general, the joint (bivariate) probability density of area-equivalent diameter and aspect ratio is not simply the product of the corresponding marginal densities, but a more sophisticated construction based on so-called copulas [8] is needed. By means of an iterative approach, we then fit bivariate probability densities  $f_{pp}^*$  and  $f_{agg}^*$  of area-equivalent diameter and aspect ratio for both primary particles and agglomerates. Subsequently, these time independent densities are used to determine the fractions  $m_t^*$  of agglomerates for each time step  $t$  of the agglomeration process. Finally, linear combinations  $(1 - m_t^*)f_{pp}^* + m_t^*f_{agg}^*$  of  $f_{pp}^*$  and  $f_{agg}^*$ , so-called mixtures, are used to model the contents of the stirred tank at each time step  $t$  of the experiments. Note that this mixture only depends on the fractions  $m_t^*$ , whereas the densities  $f_{pp}^*$  and  $f_{agg}^*$  remain constant throughout the 52 time steps, which is based on a model assumption justified in Section 2.1.

The results obtained in the present paper offer opportunities for the optimization of hydrophobic agglomeration processes. For example, for given energy inputs of the stirrer, various specifications of process parameters and feeded primary particles can be considered in order to determine their impact on the output of the agglomeration process. In particular, in future research quantitative relationships can be derived which map process parameters and model parameters of size and shape descriptors of feeded primary particles onto model parameters as well as the fraction of agglomerates obtained at the end of the agglomeration processes. A key advantage of our approach is that it allows for the determination of the initial state of agglomeration, which is crucial for characterizing the entire agglomeration process. By tracking how agglomeration evolves over time, the agglomeration process can be better understood. This insight into the temporal behavior of agglomeration processes can enable targeted adjustments to process conditions, such as stirring intensity or reactant concentrations, to achieve desired agglomerate properties. By means of these relationships, the following inverse problem can be investigated. Namely, for given (desired) model parameters of size and shape descriptors of agglomerates at the end of agglomeration processes, (optimal) values of feed and process parameters can be determined which are mapped onto the predefined model parameters of agglomerates.

The rest of the present paper is organized as follows. First, in Section 2, some preliminaries are given regarding the hydrophobic agglomeration process considered in this paper, followed by a short explanation of the methods exploited for the acquisition and processing of image data. In Section 3, the data of one of the experiments (DV1) are used as an example to explain the copula-based procedure which we apply to parametrically model the bivariate probability distributions of particle descriptor vectors. The results obtained in this paper are presented in Section 4 and discussed in Section 5. Section 6 concludes.

## 2 Some preliminaries

### 2.1 Description of the hydrophobic agglomeration process

This section briefly describes the particle system and experimental setup of the hydrophobic agglomeration process considered in [4]. Recall that agglomeration is the adherence of particles to each other after collision [9], where a system of particles can be subdivided into so-called primary particles and agglomerates. Therefore, in the following the term “particle” is used to refer to either a primary particle or an agglomerate. Note that the mentioned collision of particles requires not only a minimum level of turbulence within the liquid where the particles are situated but also a sufficient number/volume-based fraction of particles (i.e., number of particles divided by the volume of the suspension). As agglomeration reaches saturation, the number/volume-based fraction of particles decreases, causing

particles to become too sparsely distributed for further collisions. This results in a minimum number/volume-based fraction of particles.

A primary particle is a single solid particle, which means that its size and shape do not change throughout the course of the agglomeration process. An agglomerate, in contrast, is a particle consisting of multiple primary particles that adhere to each other. Since the members of a set of adhering primary particles can change during an agglomeration experiment, the size and shape of the agglomerate associated with such a set can also change over time. However, a single agglomerate’s variability in size and shape does not necessarily contradict the assumption (made in the present paper) that the agglomerates’ particle descriptors follow the same probability distribution during the entire agglomeration process. This assumption is strengthened by the circumstance that besides the growth of agglomerates, large agglomerates can be redispersed by turbulence into multiple smaller ones, which results in an equilibrium state regarding the distribution of agglomerate descriptors.

### 2.1.1 Experimental setup

In [4] the agglomeration of hydrophobic fines in a stirred tank was investigated by inline probe imaging as illustrated in Figure 1.

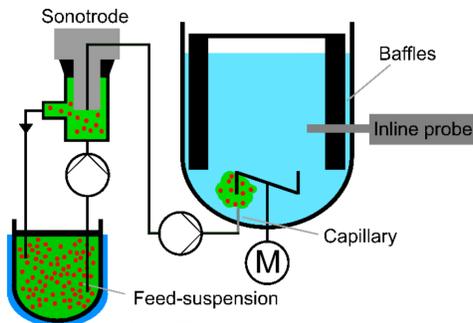


Figure 1: Experimental setup of the hydrophobic agglomeration process; taken from [4].

The experimental setup in Figure 1 shows a “feed-suspension”, which consists of silanized alumina particles and ethanol. The silanization ensures the hydrophobicity of the particles, whereas the ethanol functions as an organic solvent for dispersing the particles into almost exclusively primary particles. In Figure 2 a scanning electron microscopy (SEM) image is displayed, which highlights the almost perfectly spherical shapes of the primary particles. as well as a histogram of the number-based particle size distribution.

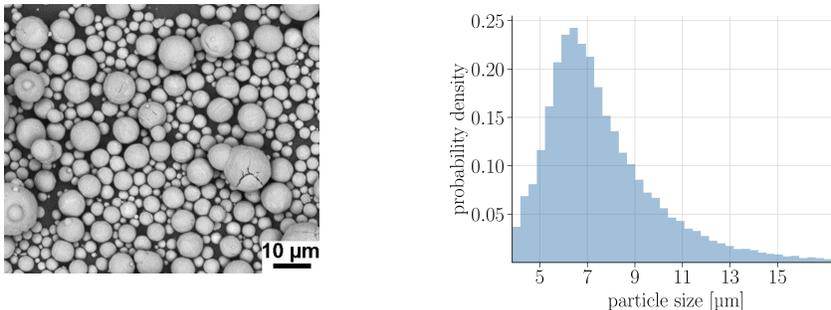


Figure 2: SEM image of silanized alumina particles (left), and histogram of the number-based particle size distribution from inline probe data for primary particles dispersed in pure ethanol (right); taken from [4].

The tank shown in Figure 1 is equipped with four baffles and is, prior to the addition of the “feed-suspension”, solely filled with 3.5 l of 0.75 mol/l NaCl solution with a pH-value adjusted to the isoelectric point of pH 7.3 for

silanized alumina particles [4]. In the experiments analyzed in the present paper, two different stirrer designs, both with 5 cm diameter, are considered.

The inclined blade (IB) stirrer uses four blades that are inclined by 45 degrees. This stirrer leads to the generation of an axial flow of the liquid from the top towards the bottom and subsequently recirculates the liquid along the walls of the tank. In contrast, the use of the Rushton-turbine, which is equipped with six upright blades, leads to the generation of axial flows from both, below and above the stirrer, and forwards them in radial directions. For reaching a constant stirrer speed of 450 rpm that is used across all experiments considered in this paper, the Rushton turbine requires a three times higher power input than the IB stirrer. As a consequence, the flow induced by the Rushton-turbine results in a three times higher (mass averaged) energy dissipation rate  $\varepsilon$  in comparison to the flow induced by the IB stirrer [10], see Table 1. Note that  $\varepsilon$  is determined by dividing the power input of the stirrer running at 450 rpm by the mass of the liquid in the tank, which is constant across all considered experiments in Table 1 that describes the three different configurations of the experiments discussed in this paper.

experiment	stirrer	$\varepsilon$ in W/kg	ultrasound
DV1/DV2	Rushton-turbine	0.3	yes
UDV1/UDV2	Rushton-turbine	0.3	no
IB3	inclined blade (IB)	0.1	yes

Table 1: Configuration of experiments, where DV2 and UDV2 are repetitions of DV1 and UDV1, respectively.

Specifically, the considered experiments are denoted by DV1, DV2, UDV1, UDV2 and IB3. Here, DV means “dispersed version”, and UDV stands for “undispersed version”, indicating whether ultrasound was applied, i.e., whether the sonotrode was running in the dispersion loop, to further disperse the particles contained in the “feed-suspension” that had already been dispersed to some extent by ethanol. By DV2 and UDV2, repetitions of DV1 and UDV1 are denoted. This means that the ratio of agglomerates to primary particles at the beginning of the experiments without ultrasonic dispersion of preexisting agglomerates, namely UDV1 and UDV2, should be higher than for the other experiments listed in Table 1.

### 2.1.2 Breakage behavior of agglomerates

The energy dissipation rate  $\varepsilon$  mentioned in Section 2.1.1 quantifies the amount of energy that is transferred from the stirrer to the liquid inside the stirred tank and thus serves as a measure for the energy available for the generation of turbulence. Therefore, the energy dissipation rate is positively correlated with the number of occurring particle collisions. Furthermore, the energy dissipation rate  $\varepsilon$  is relevant because of its influence on the Kolmogorov microscale  $\eta$  of the turbulence, which is given by

$$\eta = \left( \frac{\nu^3}{\varepsilon} \right)^{1/4}, \quad (1)$$

where  $\nu$  is the kinematic viscosity. Particles that are larger than  $\eta$  are exposed to greater shear forces that can cause agglomerates to redisperse into multiple smaller particles [2]. This means that for a larger  $\varepsilon$  smaller agglomerates can break due to shear forces. Therefore,  $\eta$  represents the approximate maximum size of agglomerates under turbulence.

## 2.2 Acquisition and processing of image data

In this section we briefly recall the methods used in [4] for the acquisition and processing of image data.

### 2.2.1 Inline camera system

Primary particles and agglomerates in the stirred tank were captured with the inline camera probe SOPAT PL (SOPAT GmbH, Germany). The probe’s immersible shaft of 12 mm diameter reaches 3.5 cm into the stirred tank. A built-in strobe allows for a good illumination of even the smallest particles. At the end of the probe a rhodium reflector is positioned at a distance of 4 mm to the camera lens. The large distance to the lens ensures that the particles in the stirred tank are sampled representatively. The camera system captures 8-bit grayscale images with a resolution of  $2464 \times 2056$  square-shaped pixels, where each pixel has a side length of  $0.2646 \mu\text{m}$ . Figure 3 (left) shows

a typical captured image, where segmented particles are outlined in red. Note that pixels that belong to particles tend to have larger grayscale values than pixels that belong to the background. In addition, Figure 3 (right) shows the distribution of grayscale values of the pixels in a total of 52,000 images captured for the experiments listed in Table 1.

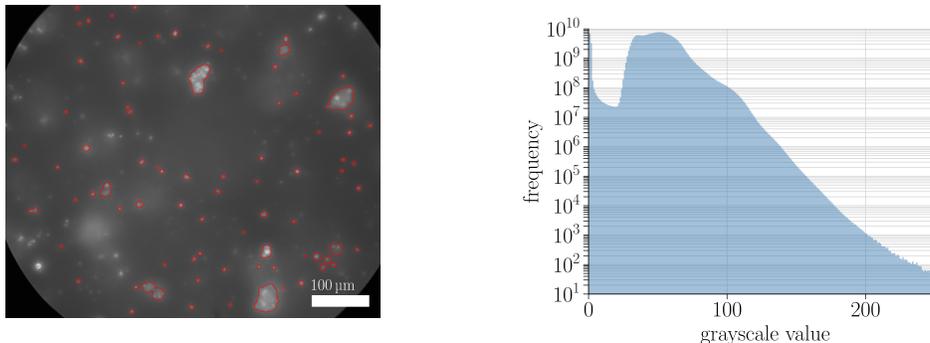


Figure 3: Contrast enhanced image (left) captured with SOPAT PL with segmented particles outlined in red, histogram (right) of grayscale values of the pixels in all images captured for the experiments listed in Table 1.

The following 47 minutes long procedure is performed for each experiment considered in this paper. First, the suspension is added to the tank with the stirrer already running. After a mixing period of 30 seconds, the image acquisition starts at time step  $t = 1$ . Then, within ten seconds, the camera system captures 200 images. After those ten seconds and an additional waiting period of 45 seconds, i.e., 55 seconds after  $t = 1$ , time step  $t = 2$  is reached. For  $t = 2$ , the same imaging procedure as for  $t = 1$  is performed. Overall, this procedure is repeated for 52 time steps, where 200 images are acquired for each time step  $t \in T = \{1, 2, \dots, 52\}$ .

### 2.2.2 Segmentation algorithm

This section summarizes the image preprocessing and segmentation steps performed in [4], where as a first step, a few undesirable effects occurring in the image data had to be removed. In some images, particles are observed that deposit directly on the lens of the probe, which appear relatively large. Moreover, these particle depositions may remain on the lens for multiple frames. Since considering these particles would lead to a non-representative particle distribution, it is desirable to skip them in the segmentation algorithm. Furthermore, there are particles that appear to be out of the focal plane. Similar as the depositions on the lens they show rather blurry edges, compare Figure 3 (left).

The initial step of the segmentation algorithm involves applying a bilateral Gaussian filter [11] to achieve image smoothing. For each pixel, this filter computes a convex combination of neighboring pixel values within a predefined window. The weights in these convex combinations depend on the distance between the pixels, taking into account both their spatial positions and grayscale value differences. The inclusion of grayscale value differences in the filter aims to preserve edges.

Even though particles appear to be brighter than the background, binarizing images with a single threshold is not suitable for reliably segmenting the particles. This is due to significantly different grayscale values of pixels associated with particles across different areas of the same image which is the result of variations in particle concentration, out-of-focus particles, and particles being attached to the lens. To address this, a range filter is applied to the image. In doing so, each pixel of the image is assigned the span of values of all neighboring pixels in a  $17 \times 17$  window around the pixel. Thus, this filter enhances particle edges while disregarding out-of-focus particles. This aims to distinguish particles in focus from out-of-focus particles and background. The resulting range-filtered image is then thresholded by setting pixels below a threshold of 9 to zero, where the pixels of unprocessed 8-bit images have grayscale values in  $\{0, 1, \dots, 255\}$ , see Figure 3. After thresholding, a morphological closing operation [12] aims to remove small connected components of non-zero pixels with 50 pixels or fewer, where the structuring element has a circular shape with a diameter of 119 pixels. The remaining connected components of non-zero pixels represent regions of interest (ROIs) for further processing.

To perform binarization, each identified ROI is processed separately. For a reliable segmentation of particles in a ROI, a combination of two binarization techniques was used: global binarization with an image specific constant threshold and a local binarization method, which computes local thresholds in the ROI based on local mean grayscale values. The results from both methods are combined by setting a pixel value to one if it exceeds at least the lower threshold of the two techniques. On the one hand, this combination is necessary due to the possibly large range in the unobservable  $z$ -direction in which larger agglomerates extend resulting in a wide range of grayscale values. For such particles, the local binarization method is inapt and a global binarization method is required. On the other hand, the global binarization method’s identification of edges of blurry objects can be unprecise. Thus, neither method is suitable to be utilized independently.

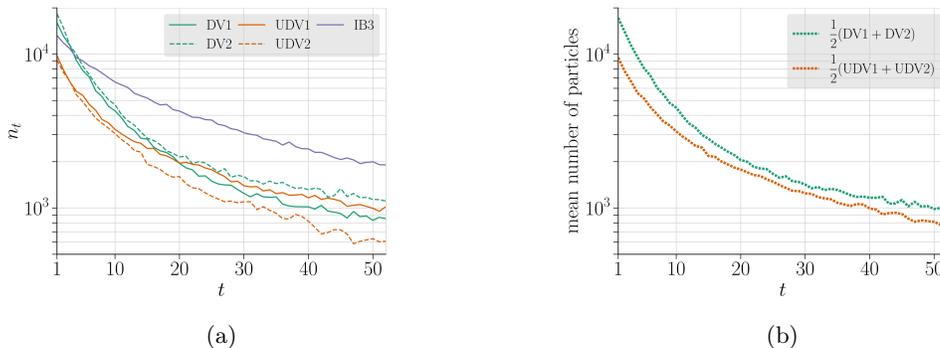


Figure 4: (a) Number  $n_t$  of particles segmented at time step  $t \in T$  for the experiments DV1, DV2, UDV1, UDV2 and IB3. (b) Mean number of particles of experiments DV1 and DV2 (green) as well as UDV1 and UDV2 (orange) at time step  $t \in T$ .

For each of the five experiments listed in Table 1, Figure 4 shows the number of particles  $n_t$  extracted from the grayscale images from the corresponding time step  $t \in T$ . Note that the decreasing number of particles in all experiments is not only caused by agglomeration where multiple primary particles are combined to a single multi-component particle. There is a small quantity of particles, mostly the larger ones, that adhere to the baffles of the tank or immobilize at the gas-liquid interface at the top of the suspension in the stirred tank, see Figure 1, where they are not captured by the inline probe.

### 3 Stochastic modeling of particle descriptor vectors

This section uses the data of experiment DV1 as an example to explain the procedure which we apply to parametrically model the probability distributions of particle descriptor vectors for each time step  $t \in T$ . First, in Section 3.1, we introduce the geometrical particle descriptors considered in this paper. In Section 3.2, we provide some fundamentals of the copula-based approach which we exploit to model the joint distribution of descriptor vectors, where we begin with modeling the marginal distributions of individual components. More precisely, we fit univariate parametric probability densities to the area-equivalent diameter and aspect ratio of the particles measured in experiment DV1 at time step  $t = 1$ , see Section 3.2.1. Then, in Section 3.2.2 the joint (bivariate) probability density of these two particle descriptors is modeled by means of so-called Archimedean copulas. The use of this class of copulas has the advantage that associated copulas typically have relatively few parameters to describe bivariate densities while offering a broad range of different copula types to flexibly capture various dependency structures. A small number of parameters has the advantage of preventing dependency structures from being overfitted, especially when only a limited amount of data is available. In comparison, non-parametric methods for modeling bivariate densities, such as kernel density estimation, can be deployed. However, since they are more flexible than copula-based approaches, they also require more data for fitting purposes—in particular to avoid overfitting [13].

The copula-based fitting of an (initial) bivariate density for descriptor vectors of agglomerates is explained in Section 3.3, whereas Section 3.4 deals with iterative adjustments of the fitted bivariate probability densities. Finally, the model validation considered in Section 3.5 provides tools for evaluating the quality of the developed model.

### 3.1 Geometrical particle descriptors

In the following, we will investigate the size and shape of particles extracted from grayscale images as described in Section 2.2. For that purpose, we consider the area-equivalent diameter  $d(P)$  and aspect ratio  $a(P)$  of the planar projection  $P \subset \mathbb{R}^2$  of three-dimensional particles.<sup>1</sup>

Recall that the area-equivalent diameter  $d(P)$  of  $P$  is given by

$$d(P) = 2\sqrt{\frac{\text{area}(P)}{\pi}}, \quad (2)$$

where  $\text{area}(P) > 0$  denotes the area of  $P$  (in  $\mu\text{m}^2$ ). Furthermore, the aspect ratio  $a(P) \in [0, 1]$  of  $P$  will be considered, which may be interpreted as a measure of roundness. In particular, an aspect ratio close to one indicates that a particle has similar width in all directions, whereas an aspect ratio close to zero corresponds to a particle with an elongated shape in one direction compared to the others. To formally define the aspect ratio  $a(P)$ , we use the Feret diameters of  $P$ , see [14]. For each pair of parallel tangents, which are touching opposite sides of a particle’s outline, the Feret diameter is given by the perpendicular distance between these tangents. The aspect ratio is then given by

$$a(P) = \frac{\text{feret}_{\min}(P)}{\text{feret}_{\max}(P)}, \quad (3)$$

where  $\text{feret}_{\min}(P)$  and  $\text{feret}_{\max}(P)$  denote the the minimum and maximum Feret diameters of  $P$ , respectively.

### 3.2 Copula-based modeling approach

As already mentioned above, the first measurement at time step  $t = 1$  shows a dominating share of primary particles. In fact, histograms of the particle descriptors  $d(P)$  and  $a(P)$  obtained from this measurement indicate unimodal probability densities, see Figure 5, whereas the presence of a larger fraction of agglomerates would lead to bi- or even multi-modality. Therefore, as an initial approach for  $t = 1$ , it seems to be reasonable to model the distributions of these descriptors by unimodal densities. However, beginning from Section 3.3 we will consider linear combinations of two unimodal densities for the descriptors of mixtures of primary particles and agglomerates.

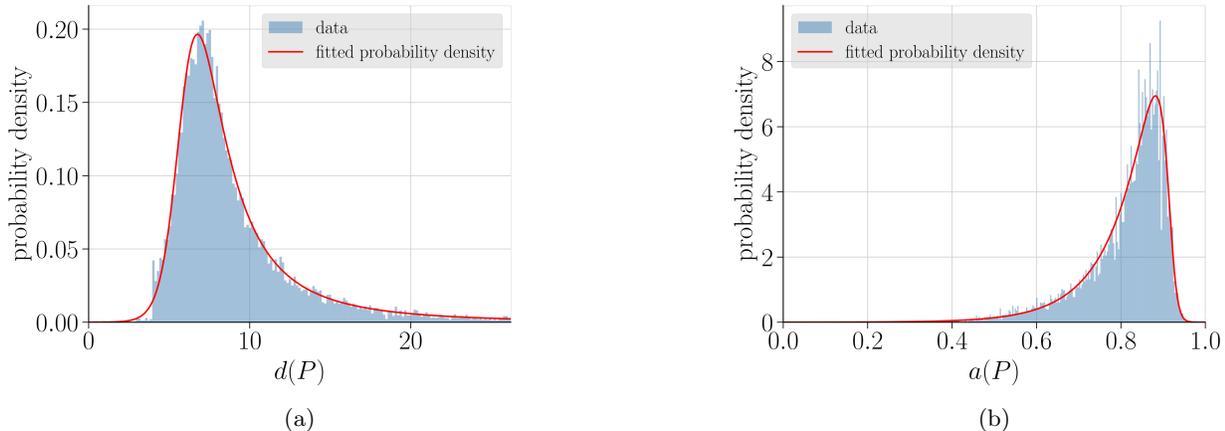


Figure 5: Parametric fits (red curves) for the univariate probability densities of area-equivalent diameter (a) and aspect ratio (b) at time step  $t = 1$  for experiment DV1.

#### 3.2.1 Univariate distribution of single particle descriptors

To get an initial model for the univariate distributions of  $d(P)$  and  $a(P)$  at time step  $t = 1$ , we consider parametric families of probability distributions from an extensive list of candidates [7] that includes, for example, the normal

<sup>1</sup>From now on, for brevity,  $P$  is called a particle, instead of saying “planar projection” of a particle.

distribution. Furthermore, in order to determine suitable parameter values of such distributions we use maximum likelihood estimation [15]. As a result, for each parametric family we get a distribution that best fits, with a corresponding likelihood. Then, we choose the parametric family with the largest likelihood as model for the distribution of the descriptor under consideration.

Among the candidate distributions provided in [7], the so-called generalized hyperbolic distribution scored the largest likelihood for both descriptors,  $d(P)$  and  $a(P)$ , where the density  $f_{\alpha,\beta,\delta,\lambda,\mu} : \mathbb{R} \rightarrow \mathbb{R}$  of the generalized hyperbolic distribution is given by

$$f_{\alpha,\beta,\delta,\lambda,\mu}(x) = \frac{e^{\beta(x-\mu)}}{\alpha^{\lambda-\frac{1}{2}}\delta\lambda\sqrt{2\pi}}(\delta^2 + (x-\mu)^2)^{\lambda-\frac{1}{2}}(\alpha^2 - \beta^2)^{\lambda/2} \frac{B_{\lambda-\frac{1}{2}}(\alpha\sqrt{\delta^2 + (x-\mu)^2})}{B_{\lambda}(\delta\sqrt{\alpha^2 - \beta^2})}, \quad (4)$$

for each  $x \in \mathbb{R}$ , see [16]. Here  $(\beta, \lambda, \mu) \in \mathbb{R}^3$ ,  $(\alpha, \delta) \in (0, \infty)^2$  with  $|\beta| < \alpha$  are model parameters. After fitting the parameters  $\alpha, \beta, \delta, \lambda, \mu$  to data and inserting them into the right-hand side of Eq. (4), the values of  $f_{\alpha,\beta,\delta,\lambda,\mu}(x)$  can be computed, see the red curves in Figure 5, where two evaluations of the modified Bessel function  $B_q$  of the third kind with index  $q$  [17] are required.

For the first time step  $t = 1$  of experiment DV1, considered in this section as an example, we obtained the following values for the model parameters  $\alpha, \beta, \delta, \lambda, \mu$ . Namely,  $\alpha = 0.7192, \beta = 0.6848, \delta = 2.2314, \lambda = -1.0307, \mu = 5.7799$  for the area-equivalent diameter  $d(P)$ , and  $\alpha = 81.5848, \beta = -72.0746, \delta = 0.0301, \lambda = 0.5085, \mu = 0.9208$  for the aspect ratio  $a(P)$ .

### 3.2.2 Bivariate distribution of two-dimensional particle descriptor vectors

If the particle descriptors  $d(P)$  and  $a(P)$  introduced in Section 3.1 could be considered to be independent random variables, the bivariate probability density  $f : \mathbb{R}^2 \rightarrow [0, \infty)$  of the two-dimensional descriptor vector  $(d(P), a(P))$  would be given by

$$f(x_1, x_2) = f_d(x_1)f_a(x_2) \quad \text{for any } x_1, x_2 \in \mathbb{R}, \quad (5)$$

where  $f_d$  and  $f_a$  denote the univariate probability densities of  $d(P)$  and  $a(P)$ , respectively. However, this approach is not suitable for correlated descriptors, as it is the case for the descriptors  $d(P)$  and  $a(P)$  of experiment DV1 at time step  $t = 1$ . Here, we obtained a Pearson correlation coefficient of  $-0.4$  for  $d(P)$  and  $a(P)$ , which is clearly distinct from zero.

One way of modeling the joint distribution of correlated particle descriptors is through multivariate normal distributions [18], with corresponding correlation matrices that capture the particle descriptors' interdependencies, like in [4]. However, the marginals of multivariate normal distributions are (univariate) normal distributions, which are symmetric and therefore inadequate to model our datasets, whose histograms are skewed, see Figure 5. As shown in Section 3.2.1, other families of parametric probability distributions, like generalized hyperbolic distributions, are more suitable for modeling the distributions of  $d(P)$  and  $a(P)$ . To come up with a parametric model for the joint distribution of the two-dimensional descriptor vector  $(d(P), a(P))$ , we consider so-called Archimedean copulas [8], which have been successfully exploited in previous studies of data for particle descriptor vectors, see e.g. [13]. Specifically, copulas allow us to incorporate both flexibility in choosing the families of marginal distributions and accurate modeling of correlations between particle descriptors into our model.

**Sklar's representation formula.** Note that the product formula given in Eq. (5) is a special case of the following representation formula for bivariate probability densities, which is a differential version of Sklar's representation formula for cumulative distribution functions, see e.g. [8]. Namely, for the bivariate density  $f : \mathbb{R}^2 \rightarrow [0, \infty)$  of the two-dimensional descriptor vector  $(d(P), a(P))$  it holds that

$$f(x_1, x_2) = c(F_d(x_1), F_a(x_2))f_d(x_1)f_a(x_2) \quad \text{for any } x_1, x_2 \in \mathbb{R}, \quad (6)$$

where  $F_d : \mathbb{R} \rightarrow [0, 1]$  and  $F_a : \mathbb{R} \rightarrow [0, 1]$  are the cumulative distribution functions of  $d(P)$  and  $a(P)$ , respectively, and  $c : [0, 1]^2 \rightarrow [0, \infty)$  is a so-called copula density, which is the bivariate probability density of a two-dimensional random vector  $(U_1, U_2)$  such that its components  $U_1$  and  $U_2$  are uniformly distributed random variables with values in the unit interval  $[0, 1]$ . In particular, in the independent case considered in Eq. (5) it holds that  $c(u_1, u_2) = 1$  for any  $u_1, u_2 \in [0, 1]$ .

**Archimedean copulas.** To parametrically model the copula density  $c : [0, 1]^2 \rightarrow [0, \infty]$  appearing in Eq. (6), we consider the cumulative distribution function  $C : [0, 1]^2 \rightarrow [0, 1]$  of  $(U_1, U_2)$  given by

$$C(u_1, u_2) = \int_0^{u_1} \int_0^{u_2} c(v_1, v_2) dv_2 dv_1 \quad \text{for any } u_1, u_2 \in [0, 1]. \quad (7)$$

Note that  $C$  is called the copula corresponding to the copula density  $c$ . In this paper we consider a special class of copulas, so-called Archimedean copulas, whose definition is based on Archimedean generators  $\varphi : [0, 1] \rightarrow [0, \infty]$ , which are continuous, strictly decreasing and convex functions with  $\varphi(0) = \infty$  and  $\varphi(1) = 0$ , see e.g. [8]. The copula  $C : [0, 1]^2 \rightarrow [0, 1]$  is then given by

$$C(u_1, u_2) = \varphi^{-1}(\varphi(u_1) + \varphi(u_2)) \quad \text{for any } u_1, u_2 \in [0, 1]. \quad (8)$$

Thus, to parametrically model the copula density  $c$  appearing in Eq. (6), we consider various parametric families  $\{\varphi_\theta : \theta \in \Theta\}$  of Archimedean generators, where  $\Theta \subset \mathbb{R}$  is some set of admissible parameters, see Table 2. Then, each family of generators  $\{\varphi_\theta : \theta \in \Theta\}$  induces a parametric family of copula densities  $\{c_\theta : \theta \in \Theta\}$ , where

$$c_\theta(u_1, u_2) = \frac{\partial^2}{\partial u_1 \partial u_2} \varphi_\theta^{-1}(\varphi_\theta(u_1) + \varphi_\theta(u_2)) \quad \text{for any } u_1, u_2 \in [0, 1] \text{ and } \theta \in \Theta. \quad (9)$$

copula	Ali-Mikhail-Haq	Clayton	Frank	Gumbel	Joe
$\Theta$	$[-1, 1)$	$(0, \infty)$	$\mathbb{R} \setminus \{0\}$	$[1, \infty)$	$[1, \infty)$
$\varphi_\theta(u)$	$\ln \frac{1-\theta(1-u)}{u}$	$\frac{1}{\theta}(u^{-\theta} - 1)$	$-\ln \frac{e^{-\theta u} - 1}{e^{-\theta} - 1}$	$(-\ln u)^\theta$	$-\ln(1-(1-u)^\theta)$

Table 2: Parametric families  $\{\varphi_\theta : \theta \in \Theta\}$  of Archimedean generators.

**Rotated Archimedean copulas.** For each of the parametric families of copula densities  $\{c_\theta : \theta \in \Theta\}$  introduced in Eq. (9), we additionally consider three rotated versions, which are obtained by rotations of  $c_\theta$  around the midpoint  $(0.5, 0.5)$  of the square  $[0, 1]^2$  by multiples of 90 degrees. In this way, we obtain a two-parametric set of copula densities  $\{c_{(\theta,r)} : \theta \in \Theta, r \in R\}$ , where  $R = \{0, 90, 180, 270\}$ . Inserting these copula densities into Eq. (6), we obtain the parametric bivariate density  $f_{(\theta,r)} : \mathbb{R}^2 \rightarrow [0, \infty)$  of the two-dimensional descriptor vector  $(d(P), a(P))$  for each  $(\theta, r) \in \Theta \times R$ , where

$$f_{(\theta,r)}(x_1, x_2) = c_{(\theta,r)}(F_d(x_1), F_a(x_2)) f_d(x_1) f_a(x_2) \quad \text{for any } x_1, x_2 \in \mathbb{R}. \quad (10)$$

Figure 6 illustrates some effects, which can occur for differently rotated Clayton copula densities  $c_{(\theta,r)}$ , where we put  $\theta = 1$ , and for the model parameters  $\alpha, \beta, \delta, \lambda, \mu$  of the (generalized hyperbolic) marginal distributions we used the values stated in Section 3.2.1.

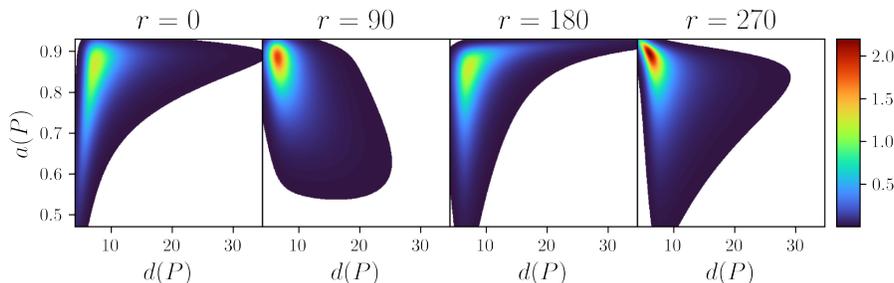


Figure 6: Bivariate density  $f_{(\theta,r)}$  of  $(d(P), a(P))$  for differently rotated Clayton copula densities  $c_{(\theta,r)}$ .

**Selection of the best fitting copula.** In Section 3.2.1, we selected the best fitting univariate marginal distributions for  $d(P)$  and  $a(P)$ , respectively, by evaluating multiple parametric distribution families using maximum likelihood estimation. Here, we follow a similar approach to select the best fitting combination of an Archimedean copula type (with its corresponding parameter  $\theta \in \Theta$ ) given in Table 2 along with a rotation angle  $r \in R$ , again based on maximum likelihood estimation. More precisely, we consider the bivariate probability density  $f_{(\theta,r)}$  given in Eq. (10), where the densities  $f_d, f_a$  (and their corresponding distribution functions  $F_d, F_a$ ) on the right-hand side have already been fitted, see Section 3.2.1. Then, the parameter  $\theta$  can be determined by means of maximum likelihood estimation, i.e., by

$$\hat{\theta} = \arg \max_{\theta \in \Theta} \mathcal{L}(\theta; D_1), \quad (11)$$

where the likelihood function  $\mathcal{L}(\theta; D_1)$  is given by

$$\mathcal{L}(\theta; D_1) = \prod_{x \in D_1} f_{(\theta,r)}(x) \quad \text{for each } \theta \in \Theta, \quad (12)$$

and  $D_1 \subset [0, \infty) \times [0, 1]$  denotes the set of all values  $x = (x_1, x_2)$  available for the descriptor vector  $(d(P), a(P))$  at time step  $t = 1$  of the experiment under consideration (i.e., DV1 in this section). Then, we choose the parametric copula type and the rotation angle, which lead to the largest likelihood  $\mathcal{L}(\hat{\theta}; D_1)$ , as modeling component for the bivariate distribution of  $(d(P), a(P))$ .

For the experiment DV1, the Clayton copula with  $\theta = 0.602$  and  $r = 90$  provided the best fit among the Archimedean copula types given in Table 2, see also Figure 7.

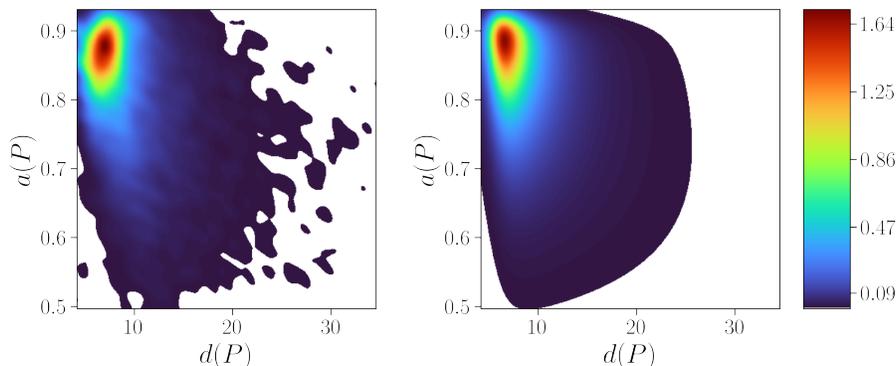


Figure 7: Bivariate probability density of  $(d(P), a(P))$  obtained for time step  $t = 1$  of experiment DV1 via kernel density estimation (left) and the parametric copula-based fitting procedure described above (right).

### 3.3 Bivariate distribution of particle descriptor vectors for agglomerates

In Section 3.2 we fitted a bivariate probability density of the two-dimensional particle descriptor vector  $(d(P), a(P))$  to data of the first measurement at time step  $t = 1$ . However, the goal of this paper is to develop a parametric model for the joint density of  $(d(P), a(P))$  for each time step  $t \in T$ , where from now on we will use the notation  $f_t$  for this density, instead of  $f$  ( $= f_1$ ) used so far. More precisely, for each  $t \in T$  we want to represent the (time-dependent) probability density  $f_t : \mathbb{R}^2 \rightarrow [0, \infty)$  of  $(d(P), a(P))$  in the following form:

$$f_t = (1 - m_t)f_{\text{pp}} + m_t f_{\text{agg}}, \quad (13)$$

where  $m_t \in [0, 1]$  is the (number-weighted) fraction of agglomerates at time step  $t \in T$ , and  $f_{\text{pp}}, f_{\text{agg}} : \mathbb{R}^2 \rightarrow [0, \infty)$  denote the bivariate probability densities of the descriptor vector  $(d(P), a(P))$  for primary particles and agglomerates, respectively, which do not depend on  $t \in T$ .

To achieve this goal, we develop an iterative approach, where the bivariate density  $f$  determined in Section 3.2 serves as initial choice for  $f_{\text{pp}}$ . It will be denoted by  $f_{\text{pp}}^{(0)}$  in the following. Furthermore, to get an initial choice for the bivariate density  $f_{\text{agg}}$  of agglomerates, denoted by  $f_{\text{agg}}^{(0)}$  from now on, we consider a dataset of values for

the descriptor vector  $(d(P), a(P))$  observed at the end of the experiment under consideration, and split it into two subsets. The reason for this is that we cannot simply assume that the particle system observed at the end of an experiment consists almost exclusively of agglomerates. Note that such a split was not necessary when determining the density  $f_{pp}^{(0)}$  ( $= f$ ), since the dataset  $D_1 \subset [0, \infty) \times [0, 1]$  considered in Section 3.3, i.e., the set of all values  $x = (x_1, x_2)$  available for  $(d(P), a(P))$  at time step  $t = 1$ , mainly consists of data pairs  $(x_1, x_2)$  corresponding to primary particles.

To reliably determine a parametric (copula-based) model for the bivariate density  $f_{agg}^{(0)}$ , we need a sufficiently large dataset with a high proportion of data pairs  $(x_1, x_2)$  corresponding to agglomerates. Therefore, we consider the union  $D_{end} = \bigcup_{t=t_{end}}^{52} D_t$  of datasets associated with several time steps close to the end of an experiment, where  $t_{end} = \min\{t \in T : n_t \leq 1.05 n_{52}\}$  and  $D_t$  denotes the set of all values  $x = (x_1, x_2)$  available for  $(d(P), a(P))$  at time step  $t \in T$ . For experiment DV1 it turned out that  $t_{end} = 46$ , which results in a number of particle descriptor vectors  $|D_{end}| = 6128$  in  $D_{end}$ , where the so-called cardinality  $|A|$  of a set  $A$  denotes the number of elements in  $A$ , i.e., in our case, the number of particle descriptor vectors. In contrast, if  $A$  is a scalar,  $|A|$  more commonly denotes the absolute value of  $A$ . As a consequence of  $D_{end}$  being the union of multiple sets  $D_t$  its cardinality  $|D_{end}|$  is much larger than the number of particles  $|D_{52}| = 853$  observed at the last time step  $t = 52$  of experiment DV1 as desired.

Moreover, to obtain a subset of  $D_{end} = \{(x_1^{(1)}, x_2^{(1)}), \dots, (x_1^{(|D_{end}|)}, x_2^{(|D_{end}|)})\}$ , which mainly consists of particle descriptor vectors corresponding to agglomerates, we determine an estimate for the (number-weighted) fraction  $m_{end} \in [0, 1]$  of agglomerates observed in  $D_{end}$ , see Figure 8.

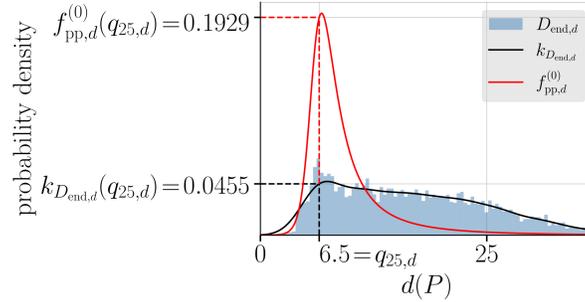


Figure 8: The values of  $q_{25,d}$ ,  $k_{D_{end},d}(q_{25,d})$  and  $f_{pp,d}^{(0)}(q_{25,d})$ , used for computing the estimate  $\hat{m}_{end}$  given in Eq. (16).

To get a formula for such an estimate, we assume that the dataset  $D_{end}$  is a sample drawn from the bivariate density  $f_{end}: \mathbb{R}^2 \rightarrow [0, \infty)$  which has the following form:

$$f_{end}(x) = (1 - m_{end})f_{pp}(x) + m_{end}f_{agg}(x) \quad \text{for each } x \in \mathbb{R}^2, \quad (14)$$

where  $f_{pp}$  and  $f_{agg}$  are the same bivariate probability densities as in Eq. (13). Furthermore, we assume that the minimum of the area-equivalent diameters of all agglomerates considered in the given experiment is larger than the 25 percent quantile  $q_{25,d}$  of the marginal density  $f_d$  representing the area-equivalent diameter in  $f_{pp}^{(0)}$ , which has been fitted in Section 3.2. This means that  $f_{agg,d}(q_{25,d}) = 0$  and, consequently, Eq. (14) gives that

$$m_{end} = 1 - \frac{f_{end,d}(q_{25,d})}{f_{pp,d}(q_{25,d})}, \quad (15)$$

where  $f_{end,d}$ ,  $f_{pp,d}$  and  $f_{agg,d}$  denote the marginal densities representing the area-equivalent diameter in  $f_{end}$ ,  $f_{pp}$  and  $f_{agg}$ , respectively. Thus, substituting the density  $f_{end,d}$  on the right-hand side of Eq. (15) by a Gaussian kernel density estimate  $k_{D_{end},d}: \mathbb{R} \rightarrow [0, \infty)$  with a bandwidth obtained by Scott's rule [19] based on the dataset  $D_{end,d} = \{x_1^{(1)}, \dots, x_1^{(|D_{end}|)}\}$ , and replacing  $f_{pp,d}$  by  $f_{pp,d}^{(0)}$ , we get the estimate

$$\hat{m}_{end} = 1 - \frac{k_{D_{end},d}(q_{25,d})}{f_{pp,d}^{(0)}(q_{25,d})}. \quad (16)$$

Note that for the data obtained in experiment DV1, it turned out that  $\widehat{m}_{\text{end}} = 0.764$ , see Figure 8.

As mentioned above, we can not assume that  $D_{\text{end}}$  consists exclusively of data corresponding to agglomerates. Thus, a fraction of  $1 - \widehat{m}_{\text{end}}$  of data corresponding to primary particles, which are encompassed within the set  $D_{\text{end}}$ , will be disregarded in order to fit the initial choice  $f_{\text{agg}}^{(0)}$  for  $f_{\text{agg}}$  to the remaining data. To be precise, we use the bivariate density  $f_{\text{pp}}^{(0)}$  to remove

$$\delta = \lceil (1 - \widehat{m}_{\text{end}}) |D_{\text{end}}| \rceil \quad (17)$$

descriptor vectors from  $D_{\text{end}}$  that are likely to correspond to primary particles, where  $\lceil x \rceil$  denotes the smallest integer which is larger than or equal to  $x > 0$ , and the remaining  $|D_{\text{end}}| - \delta$  data vectors are considered to correspond to agglomerates.

First, we draw a sample  $\Delta = \{(d_i, a_i) : i = 1, \dots, \delta\} \subset [0, \infty) \times [0, 1]$  of  $\delta$  particle descriptor vectors from a (slightly) truncated version of  $f_{\text{pp}}^{(0)}$ , using the accept-reject method [20]. Note that this method facilitates sampling a so-called target density  $f_{\text{tar}} : \mathbb{R}^2 \rightarrow [0, \infty)$ , by sampling a so-called instrumental density  $f_{\text{ins}} : \mathbb{R}^2 \rightarrow [0, \infty)$  such that  $f_{\text{tar}}$  and  $f_{\text{ins}}$  must satisfy the inequality  $f_{\text{tar}}(x) \leq M f_{\text{ins}}(x)$  for each  $x \in \{y \in \mathbb{R}^2 : f_{\text{tar}}(y) > 0\}$  of the support of  $f_{\text{tar}}$ , where  $M > 0$  is some constant.

In our case,  $f_{\text{tar}}$  is given by a truncation of  $f_{\text{pp}}^{(0)}$  whose support is restricted to the interval  $[0, q_{99.9}] \times [0, 1]$ . Here,  $q_{99.9} < \infty$  denotes the 99.9 percent quantile of the marginal probability density  $f_{\text{pp},d}^{(0)}$  of  $f_{\text{pp}}^{(0)}$ , which represents the area-equivalent diameter of primary particles. Using this truncated version of  $f_{\text{pp}}^{(0)}$ , which is a bounded function with bounded support, rules out values of area-equivalent diameter and aspect ratio outside their physically plausible ranges. Furthermore, for the instrumental density  $f_{\text{ins}}$ , we select the probability density of the uniform distribution on the rectangle  $[0, q_{99.9}] \times [0, 1]$ , putting  $M = q_{99.9} \max\{f_{\text{pp}}^{(0)}(x), x \in [0, q_{99.9}] \times [0, 1]\}$ . Then, we can sample  $f_{\text{tar}}$  by sampling  $f_{\text{ins}}$  in the following way: We draw a realization  $x \in [0, q_{99.9}] \times [0, 1]$  from  $f_{\text{ins}}$  and a realization  $u \in [0, 1]$  from the uniform distribution on the interval  $[0, 1]$ . If  $u \leq f_{\text{tar}}(x)/(M f_{\text{ins}}(x))$ , we “accept”  $x$  and, otherwise, we “reject” it. We repeat this procedure until we get “accepted” a total of  $\delta$  sample values.

The dataset  $\Delta = \{(d_i, a_i) : i = 1, \dots, \delta\} \subset [0, \infty) \times [0, 1]$  obtained in this way is now used to remove  $\delta$  descriptor vectors from  $D_{\text{end}}$ . Since the values of area-equivalent diameter and aspect ratio are on different scales, we have to normalize these quantities. For this we use the empirical standard deviations

$$\sigma_1 = \sqrt{\frac{1}{|D_{\text{end}}| - 1} \sum_{i=1}^{|D_{\text{end}}|} (x_1^{(i)} - \mu_1)^2}, \quad \sigma_2 = \sqrt{\frac{1}{|D_{\text{end}}| - 1} \sum_{i=1}^{|D_{\text{end}}|} (x_2^{(i)} - \mu_2)^2}, \quad (18)$$

where  $\mu_1 = |D_{\text{end}}|^{-1} \sum_{(d,a) \in D_{\text{end}}} d$  and  $\mu_2 = |D_{\text{end}}|^{-1} \sum_{(d,a) \in D_{\text{end}}} a$ . In the next step, we determine an injective function  $p : \{1, \dots, \delta\} \rightarrow D_{\text{end}}$  such that the descriptor vectors  $(x_1^{(p(1))}, x_2^{(p(1))}), \dots, (x_1^{(p(\delta))}, x_2^{(p(\delta))}) \in D_{\text{end}}$  can be considered to correspond to primary particles and, thus, be removed from  $D_{\text{end}}$ . To find a suitable function  $p$ , we solve the following optimization problem:

$$\min_p \sum_{i=1}^{\delta} \left\| \left( \frac{d_i - x_1^{(p(i))}}{\sigma_1}, \frac{a_i - x_2^{(p(i))}}{\sigma_2} \right) \right\| \quad (19)$$

where  $\|\cdot\|$  denotes the Euclidean norm in  $\mathbb{R}^2$ . Note that the minimization problem given in Eq. (19) can be easily reformulated as a linear assignment problem [21]. Its solution, denoted by  $p^* : \{1, \dots, \delta\} \rightarrow D_{\text{end}}$ , will be used to remove the descriptor vectors  $(x_1^{(p^*(1))}, x_2^{(p^*(1))}), \dots, (x_1^{(p^*(\delta))}, x_2^{(p^*(\delta))})$  from  $D_{\text{end}}$ , which leads to the reduced dataset

$$D_{\text{agg}} = D_{\text{end}} \setminus \left\{ (x_1^{(p^*(1))}, x_2^{(p^*(1))}), \dots, (x_1^{(p^*(\delta))}, x_2^{(p^*(\delta))}) \right\}. \quad (20)$$

Finally, using the methods stated in Section 3.2, we determine the bivariate probability density  $f_{\text{agg}}^{(0)}$  by fitting it to the dataset  $D_{\text{agg}}$  given in Eq. (20).

### 3.4 Iterative adjustment of the bivariate probability densities $f_{\text{pp}}^{(0)}$ and $f_{\text{agg}}^{(0)}$

In Sections 3.2 and 3.3 we showed how the initial fits  $f_{\text{pp}}^{(0)}$  and  $f_{\text{agg}}^{(0)}$  can be determined for the bivariate probability densities  $f_{\text{pp}}$  and  $f_{\text{agg}}$  of particle descriptor vectors of primary particles and agglomerates, respectively. However,

these fits are based on the assumption that the dataset  $D_1$ , introduced in Section 3.2, consists almost exclusively of descriptor vectors corresponding to primary particles. Thus, possible agglomerates encompassed within this dataset are erroneously considered for fitting  $f_{\text{pp}}^{(0)}$ . Furthermore, since we used  $f_{\text{pp}}^{(0)}$  to obtain the dataset  $D_{\text{agg}}$  given in Eq. (20), our assumption made on  $D_1$  also affects the fit of  $f_{\text{agg}}^{(0)}$ .

We now describe a procedure for adjusting  $f_{\text{pp}}^{(0)}$  and  $f_{\text{agg}}^{(0)}$  iteratively with the goal to successively improve their goodness-of-fit, which results in a sequence of pairs of probability densities  $(f_{\text{pp}}^{(j)}, f_{\text{agg}}^{(j)})$ , for  $j = 1, \dots, J$  with  $J \geq 2$  being the number of iteration steps in which we each adjust exactly one of the two probability densities. More precisely, in each odd iteration step  $j \in \{1, \dots, J\}$ , we determine a new probability density  $f_{\text{pp}}^{(j)}$  of primary particles, leaving the probability density  $f_{\text{agg}}^{(j-1)}$  of agglomerates unchanged. Vice versa, in each even iteration step  $j \in \{1, \dots, J\}$ , we determine a new probability density  $f_{\text{agg}}^{(j)}$  of agglomerates and leave the probability density  $f_{\text{pp}}^{(j-1)}$  of primary particles unchanged. Pseudocode for the iterative fitting procedure is provided in Algorithm 1.

---

**Algorithm 1:** Iterative adjustments and selection of best fits

---

**Input:** initially fitted densities  $f_{\text{pp}}^{(0)}$  and  $f_{\text{agg}}^{(0)}$ , sets of particle descriptor vectors  $D_1$  and  $D_{\text{end}}$ .

**Output:** final fits  $f_{\text{pp}}^*$  and  $f_{\text{agg}}^*$

· Compute  $\widehat{m}_1^{(1)}$  using Eq. (22).

· Create  $D_{\text{pp}}^{(1)}$  using Eqs. (24) and (25).

· Fit  $f_{\text{pp}}^{(1)}$  to  $D_{\text{pp}}^{(1)}$  and put  $f_{\text{agg}}^{(1)}$  equal to  $f_{\text{agg}}^{(0)}$ .

**for**  $j \in \{2, \dots, 50\}$  **do**

**for**  $z \in \{1, \text{end}\}$  **do**

        · Compute  $\widehat{m}_z^{(j)}$  and  $l_z^{(j)}$  using Eqs. (28) and (31).

**end**

**if**  $j \geq 3$  **then**

**if**  $|\widehat{m}_1^{(j)} - \widehat{m}_1^{(j-1)}| < 0.001$  and  $|\widehat{m}_{\text{end}}^{(j)} - \widehat{m}_{\text{end}}^{(j-1)}| < 0.001$  **then**

            ·  $J = j$

**break**

**end**

**end**

**if**  $j$  odd **then**

        · Compute  $\text{pp}_1(x)$  for all  $x \in D_1$  using Eq. (30).

        · Create  $D_{\text{pp}}^{(j)}$ .

        · Fit  $f_{\text{pp}}^{(j)}$  to  $D_{\text{pp}}^{(j)}$  and put  $f_{\text{agg}}^{(j)}$  equal to  $f_{\text{agg}}^{(j-1)}$ .

**end**

**else if**  $j$  even **then**

        · Compute  $\text{pp}_{\text{end}}(x)$  for all  $x \in D_{\text{end}}$  using Eq. (30).

        · Create  $D_{\text{agg}}^{(j)}$ .

        · Fit  $f_{\text{agg}}^{(j)}$  to  $D_{\text{agg}}^{(j)}$  and put  $f_{\text{pp}}^{(j)}$  equal to  $f_{\text{pp}}^{(j-1)}$ .

**end**

**end**

· Determine function  $\rho$  using Eq. (32).

·  $j^* = \arg \min_{j \in \{2, \dots, J\}} \rho(j) - 1$

·  $f_{\text{pp}}^* = f_{\text{pp}}^{(j^*)}$ ,  $f_{\text{agg}}^* = f_{\text{agg}}^{(j^*)}$

---

### 3.4.1 Algorithm for the computation of $f_{\text{pp}}^{(1)}$

As stated in Section 3.2, the probability density  $f_{\text{pp}}^{(0)}$  was fitted to the complete set  $D_1$ . In the first iteration step  $j = 1$  we now replace  $f_{\text{pp}}^{(0)}$  by  $f_{\text{pp}}^{(1)}$ , where we fit  $f_{\text{pp}}^{(1)}$  to a subset  $D_{\text{pp}}^{(1)} \subseteq D_1$  which ideally consists of the particle descriptor

vectors of all primary particles observed at time  $t = 1$ , excluding those of agglomerates. For this, we first determine an estimate of the fraction  $m_1 \in [0, 1]$  of agglomerates in  $D_1$ . Similar to Section 3.3, where we determined a 25 percent quantile, we now consider the 75 percent quantile  $q_{75,d}$  of the marginal density of  $f_{\text{agg},d}^{(0)}$ , which represents the area-equivalent diameters of agglomerates, and assume that  $q_{75,d}$  is larger than the maximum of the area-equivalent diameters of all primary particles considered in the given experiment. This means that  $f_{\text{pp},d}(q_{75,d}) = 0$  and, consequently, Eq. (13) gives that

$$m_1 = \frac{f_{1,d}(q_{75,d})}{f_{\text{agg},d}(q_{75,d})}, \quad (21)$$

where  $f_{1,d}$  denotes the marginal density representing the area-equivalent diameter in  $f_1$ . Thus, like in Section 3.3, substituting the density  $f_{1,d}$  on the right-hand side of Eq. (21) by a Gaussian kernel density estimate  $k_{D_{1,d}}: \mathbb{R} \rightarrow [0, \infty)$  based on the first components of the two-dimensional vectors in the dataset  $D_1$ , and replacing  $f_{\text{agg},d}$  by the marginal density  $f_{\text{agg},d}^{(0)}$  representing the area-equivalent diameter in  $f_{\text{agg}}$ , we get the estimate

$$\widehat{m}_1^{(1)} = \frac{k_{D_{1,d}}(q_{75,d})}{f_{\text{agg},d}^{(0)}(q_{75,d})}, \quad (22)$$

where the superscript in  $\widehat{m}_1^{(1)}$  indicates the number of the current (first) iteration step. Note that for the data obtained in experiment DV1, it turned out that  $\widehat{m}_1^{(1)} = 0.079$ , see Figure 9.

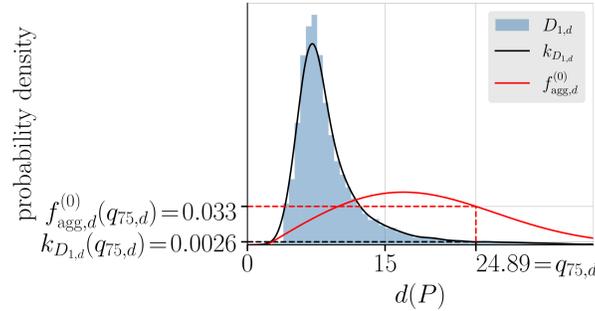


Figure 9: The values of  $q_{75,d}$ ,  $k_{D_{1,d}}(q_{75,d})$  and  $f_{\text{agg},d}^{(0)}(q_{75,d})$ , used for computing the estimate  $\widehat{m}_1^{(1)}$  given in Eq. (22).

We now show how the estimate  $\widehat{m}_1^{(1)}$  given in Eq. (22) can be used for getting a subset  $D_{\text{pp}}^{(1)} \subseteq D_1$ . In principle, we could proceed as in Section 3.3, where we determined the subset  $D_{\text{agg}} \subseteq D_{\text{end}}$  using the initial fit  $f_{\text{pp}}^{(0)}$  of  $f_{\text{pp}}$ . Analogously to this, we could exploit the initial fit  $f_{\text{agg}}^{(0)}$  of  $f_{\text{agg}}$ , derived in Section 3.3, in order to get a suitable subset  $D_{\text{pp}}^{(1)} \subseteq D_1$ . Instead, we will use another approach that includes both initial fits  $f_{\text{pp}}^{(0)}$  and  $f_{\text{agg}}^{(0)}$  of  $f_{\text{pp}}$  and  $f_{\text{agg}}$ , respectively. Using fits of the probability densities  $f_{\text{pp}}$  and  $f_{\text{agg}}$  of both particle types, primary particles and agglomerates, we expect to get a subset  $D_{\text{pp}}^{(1)} \subseteq D_1$  that is more representative for particle descriptor vectors of primary particles than just using a variant of the approach stated in Section 3.3.

More precisely, we consider the estimate  $f_1^{(1)}: \mathbb{R}^2 \rightarrow [0, \infty)$  of the bivariate probability density  $f_1$  given in Eq. (13), where

$$f_1^{(1)}(x) = (1 - \widehat{m}_1^{(1)})f_{\text{pp}}^{(0)}(x) + \widehat{m}_1^{(1)}f_{\text{agg}}^{(0)}(x) \quad \text{for each } x \in \mathbb{R}^2. \quad (23)$$

The estimate  $f_1^{(1)}$  of  $f_1$  given in Eq. (23) will be used to determine the estimate

$$\text{pp}^{(1)}(x) = (1 - \widehat{m}_1^{(1)}) \frac{f_{\text{pp}}^{(0)}(x)}{f_1^{(1)}(x)} \quad (24)$$

of the conditional probability  $\text{pp}(x) = (1 - m_1)f_{\text{pp}}(x)/f_1(x)$  that a particle observed at time step  $t = 1$  is a primary particle, under the condition that this particle has the descriptor vector  $x$ , for any given  $x \in [0, \infty) \times [0, 1]$ . This leads to the subset

$$D_{\text{pp}}^{(1)} = \left\{ x \in D_1 : \text{pp}^{(1)}(x) \geq U_x^{(1)} \right\} \quad (25)$$

of  $D_1$ , where  $\{U_x^{(1)}, x \in D_1\}$  is a family of independent random variables, which are uniformly distributed on the interval  $[0, 1]$ . Thus, by means of the methods stated in Section 3.2, we can determine the bivariate probability density  $f_{\text{pp}}^{(1)}$  by fitting it to the dataset  $D_{\text{pp}}^{(1)}$  given in Eq. (25). Afterwards,  $f_{\text{agg}}^{(1)}$  is put equal to  $f_{\text{agg}}^{(0)}$  as indicated in the beginning of Section 3.4, where we stated that we only adjust one of the two densities in each iteration step.

### 3.4.2 Procedure for subsequent adjustment of $f_{\text{pp}}^{(1)}$ and $f_{\text{agg}}^{(1)}$

Recall that we fitted the bivariate probability densities  $f_{\text{pp}}^{(1)}$  and  $f_{\text{agg}}^{(1)} (= f_{\text{agg}}^{(0)})$  to suitably chosen subsets of  $D_1$  and  $D_{\text{end}}$ , respectively. For constructing these subsets, we estimated the fractions  $m_1$  and  $m_{\text{end}}$  of agglomerates in  $D_1$  and  $D_{\text{end}}$  by considering the 25 and 75 percent quantiles  $q_{25,d}$  and  $q_{75,d}$ , where we assumed that  $f_{\text{agg},d}(q_{25,d}) = f_{\text{pp},d}(q_{75,d}) = 0$ . From now on, for each iteration step  $j \in \{2, \dots, J\}$ , we consider a different approach in order to determine estimates  $\widehat{m}_1^{(j)}$  and  $\widehat{m}_{\text{end}}^{(j)}$  of  $m_1$  and  $m_{\text{end}}$ , respectively, which will be used for constructing subsets  $D_{\text{pp}}^{(j)} \subseteq D_1$  and  $D_{\text{agg}}^{(j)} \subseteq D_{\text{end}}$ , to successively improve the fits for  $f_{\text{pp}}$  and  $f_{\text{agg}}$ . As indicated at the beginning of Section 3.4, we then fit  $f_{\text{pp}}^{(j)}$  to  $D_{\text{pp}}^{(j)}$  if  $j$  is odd, and fit  $f_{\text{agg}}^{(j)}$  to  $D_{\text{agg}}^{(j)}$  if  $j$  is even. We iterate this procedure for iteratively increasing iteration steps until the following termination condition is met for some iteration step  $j' \geq 3$ : if  $|\widehat{m}_1^{(j')} - \widehat{m}_1^{(j'-1)}| < 0.001$  and  $|\widehat{m}_{\text{end}}^{(j')} - \widehat{m}_{\text{end}}^{(j'-1)}| < 0.001$ , i.e., the estimates for  $m_1$  and  $m_{\text{end}}$  have not changed much compared to the estimates from the previous iteration, we stop iterating. In order to guarantee termination of this procedure, the procedure automatically is stopped after iteration step 50. Thus, the number of iterations  $J$  is given by

$$J = \min\{j', 50\}. \quad (26)$$

Note that, in our experiments this condition was never met, i.e., the number of iterations  $J$  was always below 50. More details on the convergence of the iterative fitting procedure can be found in Section 4.3.2 and in Figure 15.

To facilitate the explanation of the following steps, we introduce the variable  $z \in \{1, \text{end}\}$ . Specifically, for any  $j \in \{2, \dots, J\}$  and  $z \in \{1, \text{end}\}$ , we consider the likelihood function

$$\mathcal{L}_z^{(j)}(m; D_z) = \prod_{x \in D_z} \left( (1 - m)f_{\text{pp}}^{(j-1)}(x) + mf_{\text{agg}}^{(j-1)}(x) \right) \quad (27)$$

with some fraction  $m \in [0, 1]$  of agglomerates. The (maximum likelihood) estimate  $\widehat{m}_z^{(j)}$  for the fraction of agglomerates in  $D_z$  is then given by

$$\widehat{m}_z^{(j)} = \arg \max_{m \in [0, 1]} \mathcal{L}_z^{(j)}(m; D_z). \quad (28)$$

Note that in the definitions of the estimators  $\widehat{m}_{\text{end}}^{(1)} (= \widehat{m}_{\text{end}})$  and  $\widehat{m}_1^{(1)}$  given in Eqs. (16) and (22) only the area-equivalent diameters of primary particles and agglomerates are taken into account, but not their aspect ratios. However, after having determined the bivariate probability densities  $f_{\text{pp}}^{(1)}$  and  $f_{\text{agg}}^{(1)}$  in the first step, we now use the recursive definition of the likelihood function given in Eq. (27) in order to compute the estimate  $\widehat{m}_z^{(j)}$  based on descriptor vectors  $x \in D_z$  for any  $j \in \{2, \dots, J\}$  and  $z \in \{1, \text{end}\}$ , i.e., in this manner we determine an estimate  $\widehat{m}_z^{(j)}$  of the fraction  $m_z$  of agglomerates, which provides a good fit with respect to both the area-equivalent diameter and aspect ratio of particles.

Recall that in Eq. (23) we estimated  $f_1$  by  $f_1^{(1)}$  using the previously determined estimates  $f_{\text{pp}}^{(0)}$ ,  $f_{\text{agg}}^{(0)}$  and  $\widehat{m}_1^{(1)}$ . For each  $j \in \{2, \dots, J\}$ , in order to further improve the fit of the bivariate probability densities  $f_1$  and  $f_{\text{end}}$  given in Eqs. (13) and (14), we now use the estimates  $f_{\text{pp}}^{(j-1)}$ ,  $f_{\text{agg}}^{(j-1)}$  of  $f_{\text{pp}}$ ,  $f_{\text{agg}}$  obtained in the previous iteration step, as well as the estimate  $\widehat{m}_z^{(j)}$  given in Eq. (28) for  $z \in \{1, \text{end}\}$ . More precisely, for  $z \in \{1, \text{end}\}$ , we estimate  $f_z$  by  $f_z^{(j)} : \mathbb{R}^2 \rightarrow [0, \infty)$ , where

$$f_z^{(j)}(x) = (1 - \widehat{m}_z^{(j)})f_{\text{pp}}^{(j-1)}(x) + \widehat{m}_z^{(j)}f_{\text{agg}}^{(j-1)}(x) \quad \text{for each } x \in \mathbb{R}^2. \quad (29)$$

Furthermore, similar to Eq. (24), the estimate  $f_z^{(j)}$  of  $f_z$  given in Eq. (29) will be used to determine the estimate

$$\text{pp}_z^{(j)}(x) = (1 - \widehat{m}_z^{(j)}) \frac{f_{\text{pp}}^{(j-1)}(x)}{f_z^{(j)}(x)} \quad (30)$$

of the conditional probability  $\text{pp}_z(x) = (1 - m_z)f_{\text{pp}}(x)/f_z(x)$  that a particle observed in  $D_z$  is a primary particle, under the condition that it has the descriptor vector  $x$ , for any given  $x \in [0, \infty) \times [0, 1]$ . This leads to the subsets  $D_{\text{pp}}^{(j)} = \{x \in D_1: \text{pp}_1^{(j)}(x) \geq U_{1,x}^{(j)}\}$  and  $D_{\text{agg}}^{(j)} = D_{\text{end}} \setminus \{x \in D_{\text{end}}: \text{pp}_{\text{end}}^{(j)}(x) \geq U_{\text{end},x}^{(j)}\}$  of  $D_1$  and  $D_{\text{end}}$ , respectively. Here,  $\{U_{1,x}^{(j)}, x \in D_1\}$  and  $\{U_{\text{end},x}^{(j)}, x \in D_{\text{end}}\}$  are families of independent random variables, which are uniformly distributed on the interval  $[0, 1]$ . Finally, by means of the methods stated in Section 3.2, we determine either  $f_{\text{pp}}^{(j)}$  or  $f_{\text{agg}}^{(j)}$ , depending on whether  $j$  is odd or even, fitting the bivariate probability densities  $f_{\text{pp}}^{(j)}$  and  $f_{\text{agg}}^{(j)}$  to the datasets  $D_{\text{pp}}^{(j)}$  and  $D_{\text{agg}}^{(j)}$ , respectively.

### 3.4.3 Determining the best fitting pair $(f_{\text{pp}}^*, f_{\text{agg}}^*)$

In the previous section we determined a sequence of fits  $(f_{\text{pp}}^{(1)}, f_{\text{agg}}^{(1)}), \dots, (f_{\text{pp}}^{(J)}, f_{\text{agg}}^{(J)})$  for the bivariate probability densities  $f_{\text{pp}}$  and  $f_{\text{agg}}$  introduced in Eq. (13). However, in general, it is not clear if the goodness-of-fit of  $(f_{\text{pp}}^{(j)}, f_{\text{agg}}^{(j)})$  improves with increasing  $j$ . Therefore, in the following, a rule is defined which measures the suitability of  $(f_{\text{pp}}^{(j)}, f_{\text{agg}}^{(j)})$ . Recall that by means of Eqs. (27) and (28) we determined the maximum likelihood estimates  $\widehat{m}_1^{(j)}$  and  $\widehat{m}_{\text{end}}^{(j)}$  for the fractions of agglomerates in  $D_1$  and  $D_{\text{end}}$ , respectively, based on the knowledge of  $f_{\text{pp}}^{(j-1)}$  and  $f_{\text{agg}}^{(j-1)}$ . We now employ the likelihoods

$$l_z^{(j)} = \mathcal{L}_z^{(j)}(\widehat{m}_z^{(j)}; D_z), \quad \text{where } j \in \{2, \dots, J\} \text{ and } z \in \{1, \text{end}\}, \quad (31)$$

as measures for the the suitability of  $(f_{\text{pp}}^{(j)}, f_{\text{agg}}^{(j)})$  for each  $j \in \{1, \dots, J-1\}$ . More precisely, we consider the function  $\rho: \{2, \dots, J\} \rightarrow \{2, \dots, 2(J-1)\}$  which is given by

$$\rho(j) = \sum_{z \in \{1, \text{end}\}} |\{l_z^{(i)}; i \in \{2, \dots, J\}, l_z^{(i)} \geq l_z^{(j)}\}|, \quad (32)$$

for each iteration step  $j \in \{2, \dots, J\}$ . Subsequently,  $j^* = \arg \min_{j \in \{2, \dots, J\}} \rho(j) - 1 \in \{1, \dots, J-1\}$  indicates the iteration step of our final fits  $f_{\text{pp}}^* (= f_{\text{pp}}^{(j^*)})$  and  $f_{\text{agg}}^* (= f_{\text{agg}}^{(j^*)})$  for  $f_{\text{pp}}$  and  $f_{\text{agg}}$ , respectively, where the corresponding sum  $\rho(j^*)$  of ranks of the likelihoods  $l_1^{(j^*+1)}$  and  $l_{\text{end}}^{(j^*+1)}$  is the smallest.

Although we fitted  $f_{\text{pp}}^*$  and  $f_{\text{agg}}^*$  to the sets  $D_1$  and  $D_{\text{end}}$ , respectively, we assume from now on that  $f_{\text{pp}}^*$  and  $f_{\text{agg}}^*$  are well-suited for modeling the mixing components  $f_{\text{pp}}$  and  $f_{\text{agg}}$  of the bivariate probability density  $f_t$ , given in Eq. (13), for each time step  $t \in T$ .

### 3.4.4 Estimating the fraction of agglomerates $m_t$ for all $t \in T$

After determining the fits  $f_{\text{pp}}^*$  and  $f_{\text{agg}}^*$  for the bivariate probability densities  $f_{\text{pp}}$  and  $f_{\text{agg}}$  of area-equivalent diameter and aspect ratio of primary particles and agglomerates, respectively, we can now estimate the fraction of agglomerates  $m_t$  for each time step  $t \in T$ . For that purpose, we consider the probability density  $f_m^* = (1 - m)f_{\text{pp}}^* + mf_{\text{agg}}^*$  and the likelihood  $\mathcal{L}^*(m; D_t) = \prod_{x \in D_t} f_m^*(x)$  for any  $m \in [0, 1]$  and  $t \in T$ . This leads to the maximum-likelihood estimate

$$\widehat{m}_t = \arg \max_{m \in [0, 1]} \mathcal{L}^*(m; D_t) \quad (33)$$

of the fraction of agglomerates  $m_t$  for each  $t \in T$ . Furthermore, we assume that the estimates  $\{\widehat{m}_t, t \in T\}$  follow an underlying trend that is described by some parametric function  $\gamma_{a,b,\lambda}: T \rightarrow [a, b]$  such that

$$\gamma_{a,b,\lambda}(t) = a + (b - a)F_\lambda \left( \frac{t - 1}{|T| - 1} \right) \quad \text{for each } t \in T, \quad (34)$$

i.e., we assume that  $\widehat{m}_t \approx \gamma_{a,b,\lambda}(t)$  for each  $t \in T$ , with parameters  $a \in [0, 1], b \in [a, 1]$  and  $\lambda \in (0, 1)$ . Note that the input  $\frac{t-1}{|T|-1} = \frac{t-1}{52-1}$  of the function  $F_\lambda: [0, 1] \rightarrow [0, 1]$  in Eq. (34) is in  $[0, 1]$ , for each  $t \in T$ , where  $F_\lambda$  denotes the

cumulative distribution function of the continuous Bernoulli distribution [22], which is given by

$$F_\lambda(x) = \begin{cases} x, & \text{if } \lambda = 1/2, \\ \frac{\lambda^x(1-\lambda)^{1-x} + \lambda - 1}{2\lambda - 1}, & \text{otherwise,} \end{cases} \quad (35)$$

for each  $x \in [0, 1]$ . Note that the parameters  $a, b, \lambda$  in Eq. (34) are closely related with the dynamics of the agglomeration experiments considered in this paper. Specifically,  $a$  and  $b$  represent the estimated fractions  $\widehat{m}_1$  and  $\widehat{m}_{52}$  of agglomerates at the beginning and end of the experiment, i.e.,  $\widehat{m}_1 \approx \gamma_{a,b,\lambda}(1) = a$  and  $\widehat{m}_{52} \approx \gamma_{a,b,\lambda}(52) = b$ . Moreover, the function  $\gamma_{a,b,\lambda}$  given in Eq. (34) increases monotonously in  $t \in T$ , where  $\lambda < 0.5$  indicates fast agglomeration at the beginning and slower agglomeration at the end of the experiment, whereas  $\lambda > 0.5$  indicates slow agglomeration at the beginning and faster agglomeration at the end of the experiment. The closer  $\lambda$  is to 0 or 1, the more apparent the respective behaviors become. For  $\lambda = 0.5$ , the fraction of agglomerates increases linearly, see Figure 10.

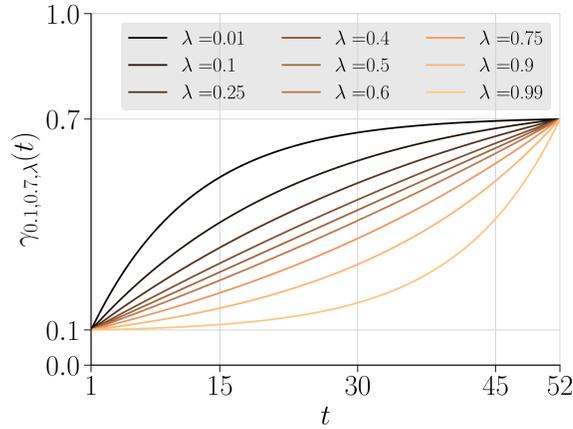


Figure 10: Visualization of  $\gamma_{a,b,\lambda}$  with  $a = 0.1$  and  $b = 0.7$ , for different values of  $\lambda \in (0, 1)$ .

Finally, we determine the (optimal) parameter vector  $(a^*, b^*, \lambda^*) \in [0, 1]^2 \times (0, 1)$  by solving the non-linear least squares problem

$$(a^*, b^*, \lambda^*) = \arg \min_{a \in [0, 1], b \in [a, 1], \lambda \in (0, 1)} \sum_{t \in T} (\widehat{m}_t - \gamma_{a,b,\lambda}(t))^2. \quad (36)$$

This gives the estimate

$$f_{m_t^*}^* = (1 - m_t^*)f_{pp}^* + m_t^*f_{agg}^*, \quad (37)$$

of the bivariate probability density  $f_t$  for each time step  $t \in T$ , where  $m_t^* = \gamma_{a^*, b^*, \lambda^*}(t)$ .

### 3.5 Model validation

We now present some methods which we use to check how well the probabilistic modeling approach developed in this paper fits the underlying image data. Recall that in the previous sections we utilized the sets  $D_1$  and  $D_{\text{end}}$  to fit the probability densities  $f_{pp}^*$  and  $f_{agg}^*$  for primary particles and agglomerates, respectively. Since we used the same kind of primary particles in all five experiments DV1, DV2, UDV1, UDV2 and IB3, a first aspect of model validation is to check if the fitted densities  $f_{pp}^*$  are similar to each other for each of these experiments, see Section 3.5.1. Furthermore, for those experiments which utilized ultrasound treatment to disperse particles into almost exclusively primary particles, the fitted values of  $a^*$  should be similar to each other, because they represent the (small) fractions of agglomerates, which are observed at the beginning of these experiments, see Section 3.5.2. Finally, we check if the probability densities  $f_{m_t^*}^*$  fit sufficiently well to their corresponding datasets  $D_t$ , for each  $t \in T$ , see Section 3.5.3. Note that for most time steps  $t \in T$ , the sets  $D_t$  were not used for fitting the components  $f_{pp}^*$  and  $f_{agg}^*$  of the mixtures  $f_{m_t^*}^* = (1 - m_t^*)f_{pp}^* + m_t^*f_{agg}^*$ .

### 3.5.1 The densities $f_{\text{pp}}^*$ and $f_{\text{agg}}^*$

Since the same type of particles is used in each of the five experiments DV1, DV2, UDV1, UDV2 and IB3, the probability densities  $f_{\text{pp}}^*$ , which represent the joint distributions of area-equivalent diameter and aspect ratio of primary particles, should be similar for all experiments listed in Table 1. On the other hand, the densities  $f_{\text{agg}}^*$  for agglomerates may depend on the specificities of the respective experimental conditions as, for example, the used energy dissipation rate generated by different types of stirrers. However, similar densities  $f_{\text{agg}}^*$  should be fitted for DV1 and DV2 (as well as for UDV1 and UDV2), since these two experiments are repetitions of each other. It turned out that in all cases a quite good fit has been achieved, see Figure 16 in Section 4.

### 3.5.2 The model parameters $a^*$ , $b^*$ and $\lambda^*$

In Section 3.4.4, see Eq. (34), we introduced the function  $\gamma_{a,b,\lambda} : T \rightarrow [a, b]$  as a parametric approximation of the estimated fractions  $\hat{m}_t$  of agglomerates for each time step  $t \in T$ , where we solved the non-linear least squares problem given in Eq. (36) in order to determine the optimal parameter values  $a^*$ ,  $b^*$  and  $\lambda^*$ . To check whether the function  $\gamma_{a,b,\lambda}$  given in Eq. (34) is an appropriate choice for a parametric model, we will investigate the mean squared error  $\text{MSE} = \frac{1}{|T|} \sum_{t \in T} (\hat{m}_t - \gamma_{a^*,b^*,\lambda^*}(t))^2$ .

Similar to the situation explained in Section 3.5.1, where we highlighted some necessary consistencies between the probability densities  $f_{\text{pp}}^*$  and  $f_{\text{agg}}^*$  for the different agglomeration experiments, some consistencies of  $a^*$ ,  $b^*$  and  $\lambda^*$  must be met. This is justified by the fact that the same type of particles is used as feed in each of the five experiments DV1, DV2, UDV1, UDV2 and IB3. Moreover, we know how the particles are inserted into the stirring tank, where in addition to ethanol, ultrasound is used (except for UDV1 and UDV2) to further disperse the “feed-suspension”, see Table 1. Therefore, the fraction of agglomerates at the beginning of the experiments, which is represented by  $a^*$ , should be similar for the experiments UDV1 and UDV2, as well as for DV1, DV2 and IB3. Furthermore, the value of  $a^*$  should be larger for UDV1 and UDV2 than for the other experiments, as there should be more agglomerates in the tank at the beginning of UDV1 and UDV2.

However, the similarities and dissimilarities of  $a^*$  mentioned above should be viewed with some caution as  $a^*$  represents the fraction of agglomerates at time step  $t = 1$ , which is not the actual beginning of an experiment. In fact, as stated in Section 2.2, the time step  $t = 1$  covers the period of 30-40 seconds after the “feed-suspension” is added to the tank. During this time period, agglomerates can already form and break, which can change the initial fraction of agglomerates. This is particularly relevant when comparing  $a^*$  for experiments DV1, DV2, and IB3, as less agglomeration in the initial 30-40 seconds is expected for IB3 since it uses a lower energy dissipation rate compared to DV1 and DV2.

Lastly, we check the similarity of the values of  $a^*$ ,  $b^*$  and  $\lambda^*$  which are obtained for the repeated experiments DV1 and DV2, as well as for UDV1 and UDV2.

### 3.5.3 The probability densities $f_{m_t^*}^*$

Finally, we check if the probability densities  $f_{m_t^*}^*$  introduced in Section 3.4.4, see Eq. (37), fit sufficiently well to their corresponding datasets  $D_t$ , for each  $t \in T$ . To accomplish this we draw a sample of 10,000 realizations from  $f_{m_t^*}^*$ , for each  $t \in T$ , utilizing the accept-reject method. In particular, to get the probability densities  $f_{\text{tar}}$  and  $f_{\text{ins}}$  required for this, we proceed in the same way as explained in Section 3.3. However, we now use  $f_{m_t^*}^*$  instead of  $f_{\text{pp}}^{(0)}$  and, as a consequence, the 99.9 percent quantile  $q_{99.9}$  is determined for the marginal probability density  $f_{m_t^*,d}^*$  of  $f_{m_t^*}^*$ , which represents the area-equivalent diameter of particles at time step  $t$ . Subsequently, for each time step  $t \in T$ , we compute various characteristics, such as mean values, standard deviations, correlation coefficients and quantiles, of the sampled two-dimensional vector data and compare them with corresponding characteristics computed for the dataset  $D_t$ .

## 4 Results

In Section 3 we presented the methods for fitting the mixture  $f_{m_t^*}^* = (1 - m_t^*)f_{\text{pp}}^* + m_t^*f_{\text{agg}}^*$  to the dataset  $D_t$ , for each time step  $t \in T$ , where we illustrated the application of these methods by means of data from experiment DV1. We now systematically present the results which we obtained for all experiments listed in Table 1.

## 4.1 Descriptive statistical analysis of measured image data

Recall that in Section 3.2, as an initial step for estimating  $f_{pp}$ , we fitted a bivariate probability density  $f_{pp}^{(0)}$  to the complete set  $D_1$ , whereas we fitted  $f_{pp}^{(j)}$  to a subset of  $D_1$ , for each  $j \in \{1, \dots, J\}$ . In fact, we initially assumed that  $D_1$  consists exclusively of primary particles. This assumption was reasonable since the univariate distributions of both the area-equivalent diameter and the aspect ratio of particles measured at time step  $t = 1$  appear to be unimodal, as shown in Figure 11. Note that this is also the case for experiments UDV1 and UDV2, where no additional ultrasound was used to disperse the particles in the “feed-suspension”.

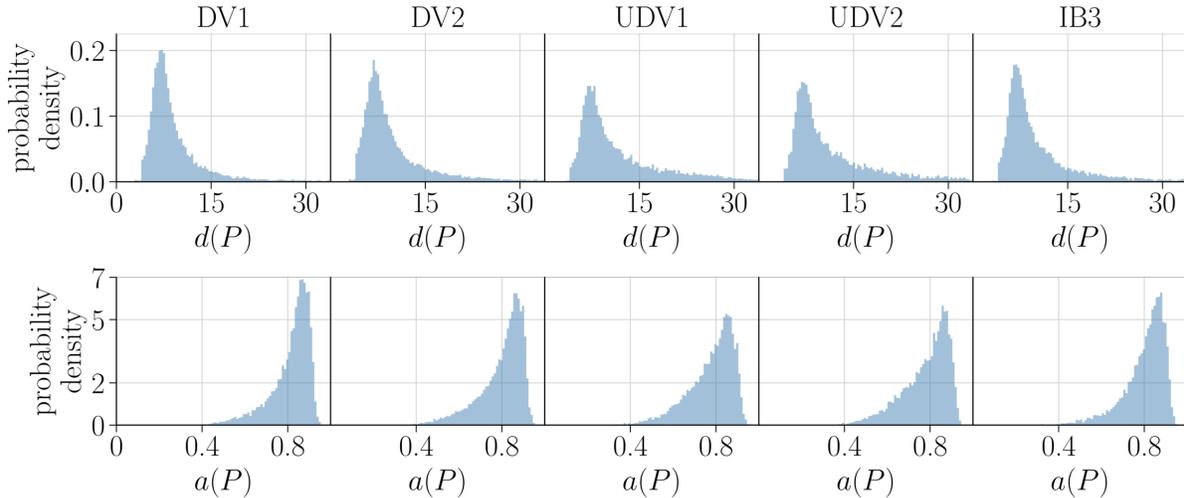


Figure 11: Histograms of area-equivalent diameter (top row) and aspect ratio (bottom row) of the descriptor vectors in  $D_1$  of the experiments DV1, DV2, UDV1, UDV2 and IB3 (from left to right).

For estimating the bivariate probability densities  $f_{pp}$  and  $f_{agg}$  we first fitted univariate parametric probability densities for the respective components, see Section 3.2.1. If the particle descriptors  $d(P)$  and  $a(P)$  represented by these univariate densities can be considered to be independent random variables with Pearson correlation coefficient equal to zero, then their joint probability density is simply given by the product of the two marginal densities, see Eq. (5). However, the empirical Pearson correlation coefficients of  $d(P)$  and  $a(P)$  shown in Figure 12 are clearly distinct from zero for all time steps  $t \in T$  and all experiments considered in this paper. Thus, we used the copula approach stated in Section 3.2.2 in order to model the joint distribution of the particle descriptors  $d(P)$  and  $a(P)$ .

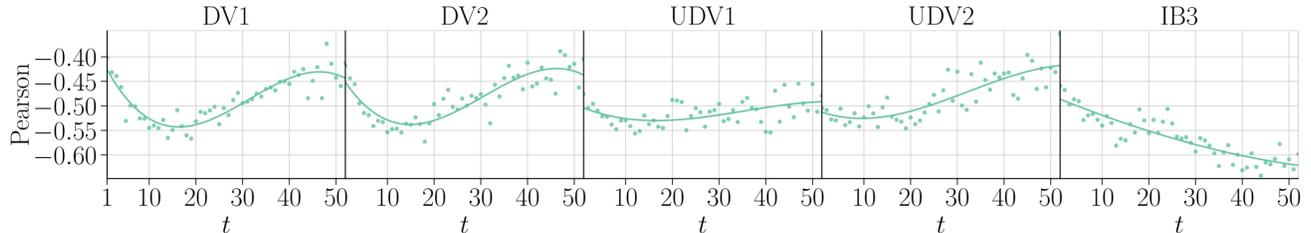


Figure 12: Empirical Pearson correlation coefficients of  $d(P)$  and  $a(P)$  for all time steps  $t \in T$  and the experiments DV1, DV2, UDV1, UDV2 and IB3. Additionally, for visualization purposes, fitted polynomials of degree 3 are shown.

## 4.2 Initial datasets associated with agglomerates

### 4.2.1 Cardinality of points in $D_{\text{end}}$

Recall that in Section 3.3, for estimating the probability density  $f_{\text{agg}}$  associated with agglomerates, we required a sufficiently large union  $D_{\text{end}} = \bigcup_{t=t_{\text{end}}}^{52} D_t$ , where  $t_{\text{end}} = \min\{t \in T : n_t \leq 1.05 n_{52}\}$ . Specifically, we expected the highest fraction of agglomerates at time steps close to the end of an experiment. The results stated in Table 3 show that the factor 1.05 in the definition of  $t_{\text{end}}$  is a reasonable choice. It leads to cardinalities  $|D_{\text{end}}|$ , i.e., numbers of particle descriptor vectors in the union sets  $D_{\text{end}}$ , which are close to the cardinalities  $|D_1|$  of  $D_1$ , see Figure 4a, and, in particular, much larger than the number of particles  $|D_{52}|$  in  $D_{52}$ , observed at the last time step  $t = 52$  of the experiments.

experiment	DV1	DV2	UDV1	UDV2	IB3
$t_{\text{end}}$	46	48	46	47	48
$ D_{\text{end}} $	6128	5719	7110	3654	9739
$ D_{52} $	853	1110	1018	608	1904

Table 3: Cardinalities of  $D_{\text{end}} = \bigcup_{t=t_{\text{end}}}^{52} D_t$  and  $D_{52}$  for the experiments DV1, DV2, UDV1, UDV2 and IB3.

### 4.2.2 Spatial arrangement of points in $D_{\text{agg}}$

In the next step, after fitting  $f_{\text{pp}}^{(0)}$  to the entire set  $D_1$ , we needed to find a subset of  $D_{\text{end}}$  that ideally consists of all descriptor vectors in  $D_{\text{end}}$  associated with agglomerates. Since no initial fit  $f_{\text{agg}}^{(0)}$  for the bivariate probability density  $f_{\text{agg}}$  of agglomerates was yet available, we had to determine such a subset  $D_{\text{agg}} \subset D_{\text{end}}$  by solving the optimization problem stated in Eq. (19). The bottom row of Figure 13 shows bivariate probability densities corresponding to the subsets  $D_{\text{agg}}$  for the experiments DV1, DV2, UDV1, UDV2 and IB3.

Note that in Figure 13 kernel density estimation has been used for visualizing the datasets  $D_{\text{end}}$ ,  $\Delta$  and  $D_{\text{agg}}$ . It turned out that in all experiments except for DV2, the largest mode of the bivariate probability density corresponding to  $D_{\text{agg}}$  is close to the largest mode of the bivariate probability density corresponding to  $\Delta$ . This is likely due to the continued presence of primary particles in  $D_{\text{agg}}$ , which becomes rather obvious in experiment IB3, where  $D_{\text{agg}}$  consists of two clearly distinct clusters of particle descriptor vectors, see Figure 13. Here, the cluster of descriptor vectors with an area-equivalent diameter of approx. 8  $\mu\text{m}$  and an aspect ratio of approx. 0.85 likely corresponds mostly to primary particles, whereas the cluster of descriptor vectors with an area-equivalent diameter of larger than 20  $\mu\text{m}$  and an aspect ratio between 0.5 and 0.85 corresponds mostly to agglomerates.

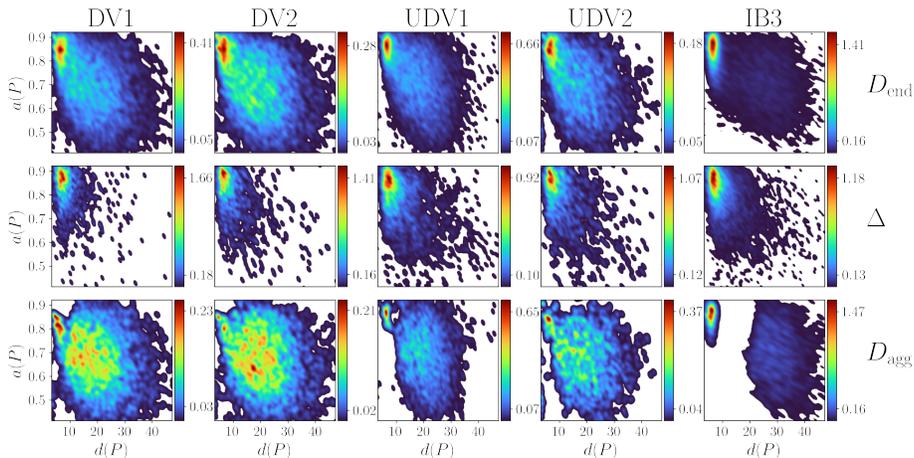


Figure 13: Bivariate probability densities corresponding to the union sets  $D_{\text{end}}$  for the experiments DV1, DV2, UDV1, UDV2 and IB3 (top row), together with densities corresponding to the sets  $\Delta$  of samples drawn from  $f_{\text{pp}}^{(0)}$  (middle row), which are used to get the subsets  $D_{\text{agg}} \subset D_{\text{end}}$  defined in Eq. (20) (bottom row).

### 4.3 Iterative adjustments and resulting fits of $f_{\text{pp}}^*$ and $f_{\text{agg}}^*$

#### 4.3.1 Spatial arrangement of points in the iteratively computed datasets $D_{\text{pp}}^{(j)}$ and $D_{\text{agg}}^{(j)}$

The spatial arrangement of data points in  $D_{\text{agg}}$ , visualized in Figure 13, suggests that in most experiments the bivariate probability density  $f_{\text{agg}}^{(0)}$  was likely fitted to a set with a non-negligible fraction of descriptor vectors associated with primary particles. However, the subsequent iterative adjustments discussed in Section 3.4 were designed to partition  $D_1$  as well as  $D_{\text{end}}$  more reliably into sets of descriptor vectors associated with either primary particles or agglomerates, requiring initial estimates  $f_{\text{pp}}^{(0)}$  and  $f_{\text{agg}}^{(0)}$  for these adjustments. Throughout the iteration steps  $j \in \{1, \dots, J\}$ , where different values of  $J$  have been chosen for each of the experiments DV1, DV2, UDV1, UDV2 and IB3 according to the rule given in Eq. (26), see Table 4, the kernel density estimates for the bivariate probability density of  $D_{\text{pp}}^{(j)}$  exhibited rather small changes.

experiment	DV1	DV2	UDV1	UDV2	IB3
$J$	4	7	14	18	30

Table 4: Number of iterations  $J$  performed for the experiments DV1, DV2, UDV1, UDV2 and IB3.

On the other hand, the experiments DV1, DV2 and IB3 showed pronounced changes in the kernel density estimates applied to  $D_{\text{agg}}^{(j)}$ . We attribute this behavior to the presumed low fraction of descriptor vectors associated with agglomerates in  $D_1$  and the higher fraction of descriptor vectors associated with primary particles in the initial subset  $D_{\text{agg}}$  of  $D_{\text{end}}$ . Thus, there is larger potential for finding subsets of  $D_{\text{end}}$  with increased representativeness of agglomerates and, accordingly, the fits  $f_{\text{agg}}^{(j)}$  also change more. As a result of this, the consecutive adjustments lead to subsets  $D_{\text{agg}}^{(j)}$  that likely contain a reduced fraction of descriptor vectors associated with primary particles as can be seen in Figure 14 for experiment IB3, where with an increasing value of  $j$ , the kernel density estimates of the bivariate probability densities corresponding to  $D_{\text{agg}}^{(j)}$  tend to have lower values for descriptor vectors  $(d(P), a(P))$  close to the mode at approximately  $(8, 0.85)$  that we discussed above.

Note that there are different scales of the color bars in Figure 14. Interestingly, the subset  $D_{\text{agg}}^{(28)}$  in the last even iteration step  $j = 28$  of experiment IB3 contains hardly any descriptor vectors close to  $(8, 0.85)$ . Moreover, for each subset  $D_{\text{agg}}^{(j)}$ , Figure 14 shows the corresponding fit  $f_{\text{agg}}^{(j)}$ , where the quality of a fit  $f_{\text{agg}}^{(j)}$  in terms of fitting the corresponding data in  $D_{\text{agg}}^{(j)}$  seems to improve in the course of iterations. For the first iterations, the main obstacle for a good fit is the bimodal distribution of the area-equivalent diameter in  $D_{\text{agg}}^{(j)}$  that is not well fitted by  $f_{\text{agg},d}^{(j)}$ . However, during these first iterations, for values of  $x = (d(P), a(P))$  close to  $(8, 0.85)$ , i.e., descriptor vectors likely associated with primary particles, the values of  $f_{\text{agg}}^{(j)}(x)$  tend to be smaller than the respective kernel density estimate of  $D_{\text{agg}}^{(j)}$  evaluated at  $x$  suggests. This, in turn, leads to a higher estimated probability  $\text{pp}_z^{(j+1)}(x)$  of a descriptor vector  $x \in D_{\text{end}}$  close to  $(8, 0.85)$  being associated with a primary particle, see Eqs. (29) and (30). Finally, this leads to the subsequent subset  $D_{\text{agg}}^{(j+1)}$  having a lower density for descriptor vectors close to  $(8, 0.85)$  than  $D_{\text{agg}}^{(j)}$ .

#### 4.3.2 Estimates $\hat{m}_1^{(j)}$ and $\hat{m}_{\text{end}}^{(j)}$ with corresponding rank sum $\rho(j)$ as measure of goodness of fit

For the construction of the subsets  $D_{\text{pp}}^{(j)}$  and  $D_{\text{agg}}^{(j)}$ , explained in Section 3.4.2, not only the bivariate probability densities  $f_{\text{pp}}^{(j)}$  and  $f_{\text{agg}}^{(j)}$  were used, but also the estimates  $\hat{m}_1^{(j)}$  and  $\hat{m}_{\text{end}}^{(j)}$  for  $m_1$  and  $m_{\text{end}}$  of the current iteration step  $j \in \{1, \dots, J\}$ . Furthermore, the stochasticity introduced by sampling from uniformly distributed random variables on the interval  $[0, 1]$  adds another layer of complexity to the computation of  $D_{\text{pp}}^{(j)}$  and  $D_{\text{agg}}^{(j)}$ .

This complexity makes it challenging to fully grasp the interplay of these factors when constructing  $D_{\text{pp}}^{(j)}$  and  $D_{\text{agg}}^{(j)}$ . However, Figure 15 attempts to shed some light on this matter, where for each experiment the estimates  $m_1^{(j)}$  and  $m_{\text{end}}^{(j)}$  as well as the rank sum  $\rho(j)$  in each iteration step  $j \in \{2, \dots, J\}$  is shown. Moreover, the mixing ratios  $\hat{m}_1^{(1)}$  and  $\hat{m}_{\text{end}}^{(1)}$  estimated by means of quantiles are included in gray boxes, whereas the mixing ratios  $m_1^{(j)}$  and  $m_{\text{end}}^{(j)}$ , which are obtained by maximum-likelihood estimation as stated in Eq. (28), are visualized on the white background.

Recall that the number of iterations  $J$  is always less than 50, see Eq. (26). Furthermore, in each experiment the termination condition introduced in Section 3.4.2 is met. This means that at some iteration step the adjustment

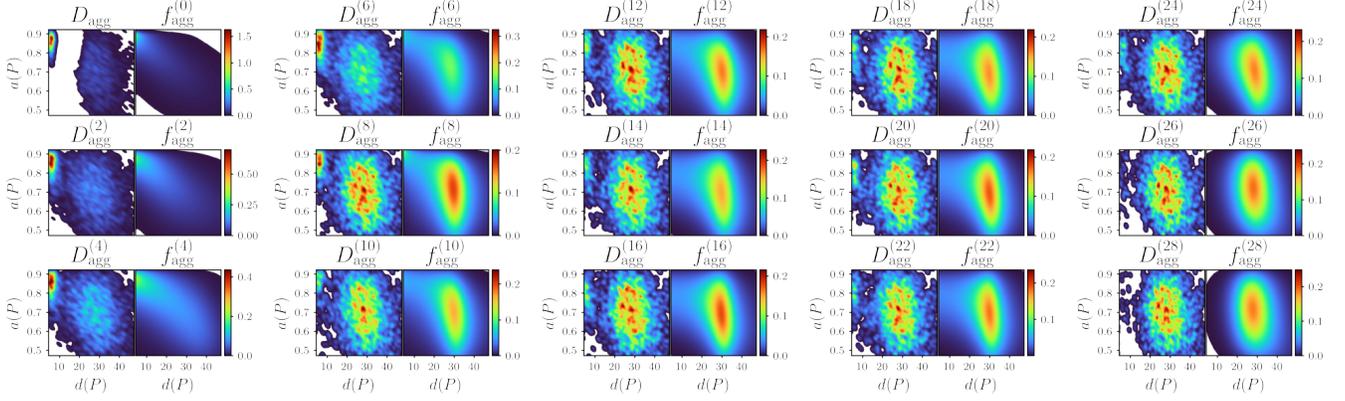


Figure 14: Kernel density estimate of  $D_{\text{agg}}$  and corresponding bivariate probability density  $f_{\text{agg}}^{(0)}$  for experiment IB3 (top left), followed by the result obtained for  $D_{\text{agg}}$  and  $f_{\text{agg}}^{(j)}$  for all even iteration steps  $j \in \{2, \dots, 28\}$ .

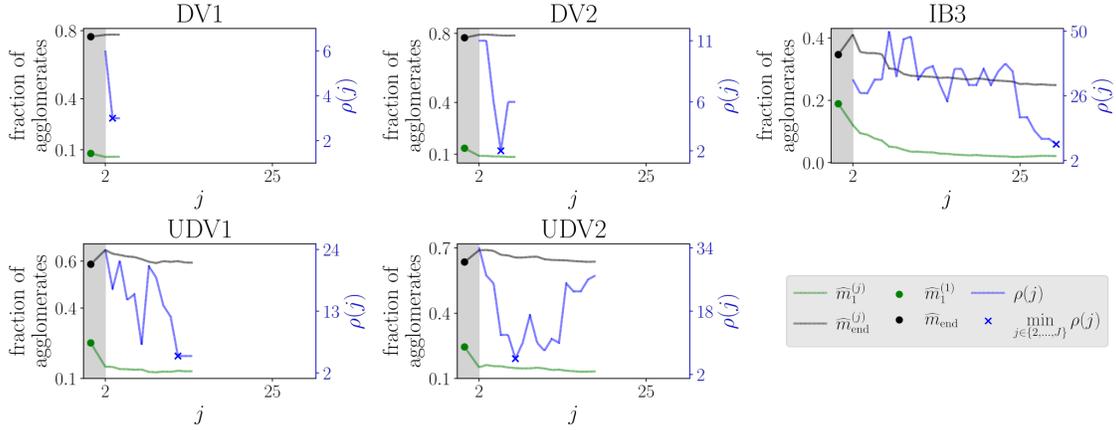


Figure 15: Estimates  $\widehat{m}_1^{(1)}$  and  $\widehat{m}_{\text{end}}$  (dots), together with subsequent estimates  $\widehat{m}_1^{(j)}$  (green) and  $\widehat{m}_{\text{end}}^{(j)}$  (black) of the fraction of agglomerates in  $D_1$  and  $D_{\text{end}}$ , for each iteration step  $j \in \{2, \dots, J\}$ . The corresponding rank sum  $\rho(j)$  is shown in blue, where an additional y-axis on the right-hand side of each plot is given.

of the estimates  $\widehat{m}_1^{(j-1)}$  and  $\widehat{m}_{\text{end}}^{(j-1)}$  from the previous iteration becomes marginal. For experiments DV1 and DV2 this happens at early iteration steps  $j = 4$  and  $j = 7$ , respectively. In all iteration steps of both experiments, the estimates for the mixing ratios hardly change, both in absolute and relative terms. On the other hand, the numbers of adjustments  $J$  in experiments UDV1 and UDV2 are higher with  $J = 14$  and  $J = 18$ , respectively. Moreover, in comparison to experiments DV1 and DV2, the mixing ratios change only slightly more. However, the adjustments of  $\widehat{m}_{\text{end}}$  to  $\widehat{m}_{\text{end}}^{(2)}$  and  $\widehat{m}_1^{(1)}$  to  $\widehat{m}_1^{(2)}$  are quite considerable. Finally, in experiment IB3, the number of adjustments  $J = 30$  is the highest and the changes of mixing ratios are most significant in the first iteration steps, whereas the changes become smaller in later iterations. Overall, the large differences between  $\widehat{m}_{\text{end}}$  and  $\widehat{m}_1^{(j^*)}$  or between  $\widehat{m}_1^{(1)}$  and  $\widehat{m}_1^{(j^*)}$ , occurring in some cases, highlight the necessity of these iterative adjustments.

Since neither the mixing ratios  $\widehat{m}_1^{(j)}$  and  $\widehat{m}_{\text{end}}^{(j)}$  nor the rank sum  $\rho(j)$  behave monotonously in  $j$ , it is difficult to evaluate the quality of each of the changes. In experiment UDV2, for example, after iteration  $\arg \min_{j \in \{2, \dots, J\}} \rho(j) = 7$  many iterations follow, in which  $\rho(j)$  tends to increase. Although, as we discussed above in the context of Figure 14, the subsets  $D_{\text{agg}}$  in experiment IB3 seem to drastically improve over the course of iterations, the rank sum  $\rho(j)$  is clearly not monotonously decreasing. Therefore, we directly check whether  $j^*$ , i.e., the argument of the minimum of  $\rho(j)$  minus 1, was nonetheless a good choice for determining  $f_{\text{pp}}^* (= f_{\text{pp}}^{(j^*)})$  and  $f_{\text{agg}}^* (= f_{\text{agg}}^{(j^*)})$ .

### 4.3.3 Final parametric fits $f_{pp}^*$ and $f_{agg}^*$ for primary particles and agglomerates

In Figure 16, the bivariate probability densities  $f_{pp}^*$  and  $f_{agg}^*$  are illustrated for each experiment given in Table 1. As detailed in Section 3.5.1, these densities are subject to certain consistency conditions across the experiments DV1, DV2, UDV1, UDV2 and IB3, which we can now verify. First, the probability densities  $f_{pp}^*$  across these experiments appear notably similar, confirming our expectation based on the use of the same particle type in each experiment. However, while the modes of all densities  $f_{pp}^*$  are located around  $(8, 0.85)$ , the value of  $f_{pp}^*$  evaluated at its mode is significantly higher for experiment DV1. At the same time, the values of  $f_{pp}^*$  for experiment DV1 are smaller for descriptor vectors with a large area-equivalent diameter and a small aspect ratio. Nevertheless, overall the probability densities appear to be very similar as desired. Moreover, the probability densities  $f_{agg}^*$  for experiments DV1 and DV2 are almost identical, as expected, given that these experiments are repetitions of each other. However, for the other pair of repeated experiments, UDV1 and UDV2, while the modes of their respective fits  $f_{agg}^*$  are located similarly, the variance of the area-equivalent diameter is smaller and the variance of the aspect ratio is larger for  $f_{agg}^*$  of UDV1, compared to  $f_{agg}^*$  of UDV2. Notably, the bivariate density  $f_{agg}^*$  of IB3 is most different from the fits of the other experiments, since it exhibits agglomerates that are larger.

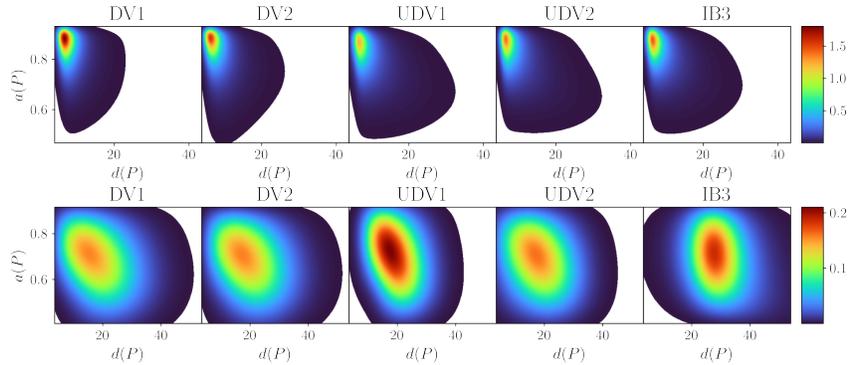


Figure 16: Visualization of the fitted bivariate probability densities  $f_{pp}^*$  (top row) and  $f_{agg}^*$  (bottom row) for the experiments DV1, DV2, UDV1, UDV2 and IB3.

For the bivariate probability densities depicted in Figure 16, the parametric families of both, the two marginal densities of area-equivalent diameter and aspect ratio, and the copula, along with the values of their respective fitted parameters, are listed in Table 5. Except for the probability density of the generalized hyperbolic distribution, which is given in Eq. (4), the probability densities of the other parametric families of univariate distributions appearing in Table 5 are given in Table 6.

## 4.4 Model parameters for the time-dependent fraction of agglomerates

In Section 3.4.4 we fitted a parametric function  $\gamma_{a,b,\lambda} : T \rightarrow [a, b]$ , see Eq. (34), to the estimated fractions  $\hat{m}_t$  of agglomerates for each time step  $t \in T$  by solving the non-linear least squares problem stated in Eq. (36). The small values of MSE given in Table 7 show that the parametric function  $\gamma_{a,b,\lambda} : T \rightarrow [a, b]$  introduced in Eq. (34) fits the estimated fractions  $\hat{m}_t$  of agglomerates for all  $t \in T$  quite well.

Recall that the optimal parameter values  $a^*, b^*, \lambda^*$  provided in Table 7 are closely related with the dynamics of the agglomeration experiments. For example,  $a^*$  and  $b^*$  represent the estimated fractions  $\hat{m}_1$  and  $\hat{m}_{52}$  of agglomerates at the beginning and end of the experiment, i.e.,  $\hat{m}_1 \approx \gamma_{a^*, b^*, \lambda^*}(1) = a^*$  and  $\hat{m}_{52} \approx \gamma_{a^*, b^*, \lambda^*}(52) = b^*$ . Moreover, in Section 3.5.2 we discussed some consistencies between  $a^*, b^*$  and  $\lambda^*$  which we verify in the following. Namely, the fraction of agglomerates at the beginning of the experiments, represented by  $a^*$ , should be similar for the experiments UDV1 and UDV2, as well as for DV1, DV2 and IB3. Furthermore, the value of  $a^*$  should be larger for UDV1 and UDV2 than for the remaining experiments. Remarkably, the mean value of  $a^*$  for experiments DV1, DV2 and IB3 is only slightly above 0 (with a very small mean agglomerate fraction of 1.23%), whereas the mean value of  $a^*$  for experiments UDV1 and UDV2, in which no ultrasound was used, is significantly higher (with a mean agglomerate fraction of 11.35%). The use or non-use of ultrasound is the only difference between the pairs of repeated experiments DV1, DV2 and UDV1, UDV2, which not only leads to lower values of  $a^*$  in DV1 and DV2, but also to higher values of  $b^*$  than in the experiments UDV1 and UDV2.

density/ experiment	particle descriptor	marg. distrib./ copula type	fitted parameter values
$f_{pp}^*$ , DV1	area-eq. dia.	gen. hyperbolic	$\alpha = 0.9047, \beta = 0.8932, \delta = 2.7427, \lambda = -1.8455, \mu = 5.4743$
	aspect ratio	gen. hyperbolic	$\alpha = 93.8323, \beta = -84.3239, \delta = 0.0373, \lambda = 0.0805, \mu = 0.9255$
		Clayton	$\theta = 0.4894, r = 90$
$f_{pp}^*$ , DV2	area-eq. dia.	gen. hyperbolic	$\alpha = 1.0632, \beta = 1.0439, \delta = 2.7612, \lambda = -1.6566, \mu = 4.9977$
	aspect ratio	gen. hyperbolic	$\alpha = 71.3184, \beta = -60.9841, \delta = 0.0254, \lambda = 0.8117, \mu = 0.9154$
		Ali-Mikhail-Haq	$\theta = 0.8626, r = 90$
$f_{pp}^*$ , UDV1	area-eq. dia.	gen. hyperbolic	$\alpha = 0.6736, \beta = 0.6212, \delta = 1.8247, \lambda = -0.3727, \mu = 6.0196$
	aspect ratio	exp. Weibull	$\alpha = 39.0851, \theta = 0.2351, \sigma = 1.0722, \mu = -0.1818$
		Ali-Mikhail-Haq	$\theta = 0.9154, r = 90$
$f_{pp}^*$ , UDV2	area-eq. dia.	gen. hyperbolic	$\alpha = 0.8722, \beta = 0.8147, \delta = 1.5869, \lambda = -0.3266, \mu = 5.7841$
	aspect ratio	exp. Weibull	$\alpha = 44.1723, \theta = 0.1915, \sigma = 0.9934, \mu = -0.0934$
		Ali-Mikhail-Haq	$\theta = 0.9458, r = 90$
$f_{pp}^*$ , IB3	area-eq. dia.	gen. hyperbolic	$\alpha = 0.8009, \beta = 0.7421, \delta = 1.7527, \lambda = -0.4319, \mu = 5.6903$
	aspect ratio	exp. Weibull	$\alpha = 99.3528, \theta = 0.2888, \sigma = 2.7903, \mu = -1.8971$
		Ali-Mikhail-Haq	$\theta = 0.9249, r = 90$
$f_{agg}^*$ , DV1	area-eq. dia.	Burr Type XII	$c = 2.2112, k = 30.0675, \sigma = 90.1895, \mu = 1.9944$
	aspect ratio	beta	$\alpha = 4.4876, \beta = 3.0262, \sigma = 0.6849, \mu = 0.2693$
		Ali-Mikhail-Haq	$\theta = -0.8795, r = 180$
$f_{agg}^*$ , DV2	area-eq. dia.	Burr Type XII	$c = 2.4661, k = 14.7549, \sigma = 60.9572, \mu = 1.9904$
	aspect ratio	beta	$\alpha = 4.3865, \beta = 2.9258, \sigma = 0.6719, \mu = 0.2724$
		Ali-Mikhail-Haq	$\theta = -0.7788, r = 180$
$f_{agg}^*$ , UDV1	area-eq. dia.	power normal	$\alpha = 0.1236, \sigma = 2.924, \mu = 11.5114$
	aspect ratio	beta	$\alpha = 3.7349, \beta = 2.7293, \sigma = 0.6281, \mu = 0.3343$
		Ali-Mikhail-Haq	$\theta = -0.9584, r = 180$
$f_{agg}^*$ , UDV2	area-eq. dia.	rice	$\lambda = 1.6661, \sigma = 8.7099, \mu = 1.0592$
	aspect ratio	Johnson S <sub>b</sub>	$\gamma = -0.3208, \delta = 1.3273, \sigma = 0.6877, \mu = 0.3009$
		Frank	$\theta = -1.1435, r = 0$
$f_{agg}^*$ , IB3	area-eq. dia.	Johnson S <sub>U</sub>	$\alpha = -0.5473, \beta = 1.7638, \sigma = 11.6947, \mu = 25.2372$
	aspect ratio	beta	$\alpha = 2.9333, \beta = 2.3599, \sigma = 0.5718, \mu = 0.372$
		Clayton	$\theta = 0.2155, r = 90$

Table 5: Parametric families of marginal distributions and copula types of  $f_{pp}^*$  and  $f_{agg}^*$ , for the experiments DV1, DV2, UDV1, UDV2 and IB3, together with the fitted parameter values (gen. = generalized, exp. = exponentiated).

parametric family	probability density	support	parameters
exp. Weibull [7, 23]	$\sigma^{-1} \alpha \theta (1 - e^{-\frac{x-\mu}{\sigma}})^{\theta-1} e^{-\frac{(x-\mu)^\alpha}{\sigma}} \frac{(x-\mu)^{\alpha-1}}{\sigma}$	$[\mu, \infty)$	$\mu \in \mathbb{R}$ and $\alpha, \theta, \sigma \in (0, \infty)$
Burr Type XII [7, 24]	$\sigma^{-1} c k \frac{(x-\mu)^{c-1}}{\sigma} (1 + \frac{(x-\mu)^c}{\sigma})^{-(k+1)}$	$[\mu, \infty)$	$\mu \in \mathbb{R}$ and $c, k, \sigma \in (0, \infty)$
beta [7, 25]	$\frac{(\frac{x-\mu}{\sigma})^{\alpha-1} (1 - \frac{x-\mu}{\sigma})^{\beta-1}}{\sigma B(\alpha, \beta)}$	$[\mu, \mu + 1]$	$\mu \in \mathbb{R}$ and $\alpha, \beta, \sigma \in (0, \infty)$
power normal [7, 26]	$\sigma^{-1} \alpha \phi(\frac{x-\mu}{\sigma}) \Phi^{\alpha-1}(-\frac{x-\mu}{\sigma})$	$\mathbb{R}$	$\mu \in \mathbb{R}$ and $\alpha, \sigma \in (0, \infty)$
Rice [7, 27]	$\frac{x-\mu}{\sigma^2} e^{-\frac{1}{2}((x-\mu)^2/\sigma^2 + \lambda^2)} B_0(\frac{x-\mu}{\sigma}, \lambda)$	$(\mu, \infty)$	$\mu \in \mathbb{R}$ and $\lambda, \sigma \in (0, \infty)$
Johnson S <sub>b</sub> [7, 28]	$\frac{\delta \sigma}{(x-\mu)(\sigma-x+\mu)} \phi(\gamma + \delta \log \frac{x-\mu}{\sigma-x+\mu})$	$(\mu, \mu + 1)$	$\gamma, \mu \in \mathbb{R}$ and $\delta, \sigma \in (0, \infty)$
Johnson S <sub>U</sub> [7, 29]	$\frac{\beta}{\sqrt{(x-\mu)^2 + \sigma^2}} \phi(\alpha + \beta \sinh^{-1} \frac{x-\mu}{\sigma})$	$\mathbb{R}$	$\alpha, \mu \in \mathbb{R}$ and $\beta, \sigma \in (0, \infty)$

Table 6: Parametric families of univariate distributions with corresponding density, support and parameter space. The symbols B, B<sub>0</sub>, sinh<sup>-1</sup> denote the beta function [17], modified Bessel function of the first kind with index 0 [17] and the inverse of the hyperbolic sine function, whereas  $\phi$  and  $\Phi$  denote the probability density and cumulative distribution function of the standard normal distribution.

As stated above, we checked the similarity of the values of  $a^*$ ,  $b^*$  and  $\lambda^*$  for the repeated experiments DV1 and DV2, as well as for UDV1 and UDV2. For DV1 and DV2 these values are quite similar to each other, whereas the

experiment	$a^*$	$b^*$	$\lambda^*$	MSE
DV1	0.0129	0.8074	0.0810	$5.217 \cdot 10^{-4}$
DV2	0.0163	0.8217	0.1105	$7.211 \cdot 10^{-4}$
UDV1	0.1027	0.5978	0.0146	$3.029 \cdot 10^{-4}$
UDV2	0.1243	0.6749	0.1190	$3.820 \cdot 10^{-4}$
IB3	0.0078	0.2543	0.3013	$8.091 \cdot 10^{-5}$

Table 7: Optimal parameter values  $a^*$ ,  $b^*$  and  $\lambda^*$  of the function  $\gamma_{a,b,\lambda} : T \rightarrow [a, b]$  introduced in Eq. (34), together with the corresponding MSE, for the experiments DV1, DV2, UDV1, UDV2 and IB3.

values of  $a^*$ ,  $b^*$  and  $\lambda^*$  behave differently for UDV1 and UDV2, see Table 7. While the difference between the two values of  $a^*$  is also small for UDV1 and UDV2, the differences for  $b^*$  and  $\lambda^*$  are larger. Recall that in Figure 10 we discussed the influence of the parameter  $\lambda$  on the shape of the function  $\gamma_{a,b,\lambda} : T \rightarrow [a, b]$ . In particular, the value of  $\lambda^* = 0.1190 \approx 0.1$  obtained for UDV2 indicates a slower increase of the fraction of agglomerates at the beginning and a stronger increase towards the end of the agglomeration experiment compared to the value of  $\lambda^* = 0.0146 \approx 0.01$  obtained for UDV1. Finally, note that the largest value of  $\lambda^* = 0.3013$  has been obtained for experiment IB3, which differs clearly from the values of  $\lambda^*$  for the remaining experiments. For  $\lambda^* = 0.3013$ , there is still a non-negligible gradient at  $t = 52$ , in contrast to the situation observed for the remaining experiments at a higher energy dissipation rate, see Figure 10. This suggests the possibility that the increase in the fraction of agglomerates has not yet reached saturation at the end of experiment IB3, which is a consequence of the lower energy dissipation rate for IB3, and the therewith associated lower probability of collision between the particles.

#### 4.5 Distribution of the time-dependent content in the stirred tank

For each time step  $t \in T$ , we modeled the bivariate probability density  $f_{m_t}^*$  of particle descriptor vectors in  $D_t$  as a mixture of  $f_{pp}^*$  and  $f_{agg}^*$ , see Eq. (37) in Section 3.4.4. Figure 17 shows the fitted density  $f_{m_t}^*$  of experiment DV1 at time steps  $t \in \{1, 26, 52\}$  or in other words, approximately 0.5, 23.5 and 47 minutes after the “feed-suspension” is added to the stirred tank.

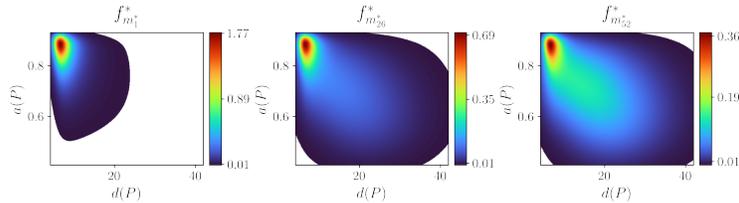


Figure 17: Fitted bivariate probability densities  $f_{m_1}^*$  (left),  $f_{m_{26}}^*$  (middle) and  $f_{m_{52}}^*$  (right) of experiment DV1.

So far, we evaluated the goodness-of-fit of the components  $f_{pp}^*$ ,  $f_{agg}^*$  and  $m_{t^*}^* (= \gamma_{a^*,b^*,\lambda^*}(t))$  of the mixture  $f_{m_t}^*$ . However, in general, we can not conclude from this, whether  $f_{m_t}^*$  itself fits sufficiently well to its corresponding dataset  $D_t$ . Therefore, as explained in Section 3.5.3, we generate (artificial) two-dimensional vector data sampled from  $f_{m_t}^*$  and compare various characteristics of them with the corresponding characteristics computed for the dataset  $D_t$ . First, we examine the empirical correlation coefficient between area-equivalent diameter and aspect ratio. For experiments DV1, DV2, and UDV2, the values of this characteristic computed from measured and simulated data, respectively, are quite close to each other, while the differences for experiments UDV1 and IB3 are somewhat larger, see Figure 18.

In Figure 19, mean value functions along with corresponding standard deviation bands are visualized for area-equivalent diameter and aspect ratio across all time steps  $t \in T$  for the experiments DV1, DV2, UDV1, UDV2 and IB3, where the purple areas highlight the intersections of the standard deviation bands computed for values of the respective descriptor in  $D_t$  (green) and in the simulated dataset drawn from  $f_{m_t}^*$  (orange). Remarkably, the curves representing the mean values are nearly identical across all instances. Moreover, the standard deviation bands for the aspect ratio computed for the measured datasets  $D_t$  and the simulated data drawn from  $f_{m_t}^*$ , respectively, coincide nicely, whereas the standard deviation bands for the area-equivalent diameter differ slightly from each other in some

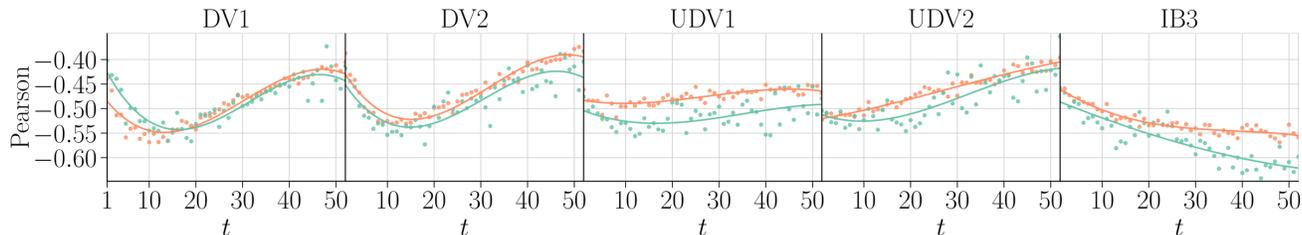


Figure 18: Empirical Pearson correlation coefficients of the particle descriptors  $d(P)$  and  $a(P)$  in  $D_t$  (green dots) and in samples of 10,000 realizations drawn from  $f_{m_t}^*$  (orange dots), for each  $t \in T$ . Additionally, like in Figure 12, fitted polynomials of degree 3 are shown in the respective colors.

cases. However, overall Figure 19 shows a very good match between the mean value functions and standard deviation bands computed for measured and simulated values of particle descriptors, respectively.

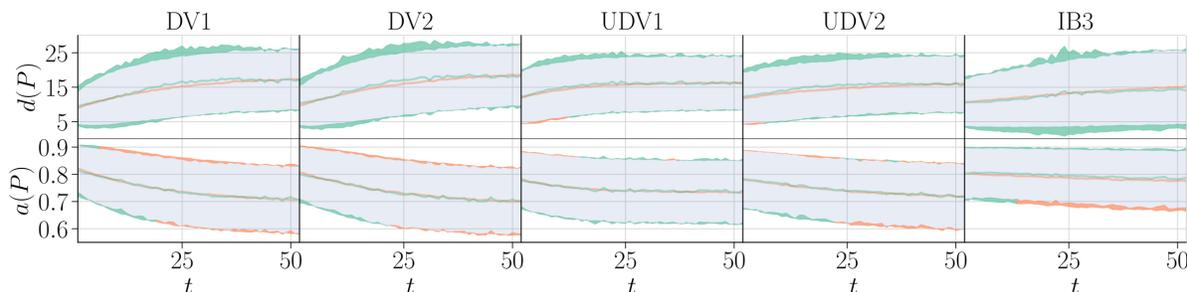


Figure 19: Time-dependent statistics of area-equivalent diameter (top row) and aspect-ratio (bottom row). Specifically, the dashed lines represent the means of the respective descriptor in  $D_t$  (green) and in the 10,000 realizations drawn from  $f_{m_t}^*$  (orange), for each time step  $t \in T$ . Corresponding standard deviation bands are plotted in the respective colors, where intersections of green and orange bands are visualized in purple.

Finally, we investigate the 5 percent, 50 percent (median), and 95 percent quantiles of the values of area-equivalent diameter and aspect-ratio computed from  $D_t$  and simulated data, respectively, for each time step  $t \in T$ . The obtained results are shown in Figure 20. Here, one can see that the quantiles of aspect ratio computed from from measured and simulated data are almost identical for each time step  $t \in T$  and for each experiment. In addition, 5 percent quantile and median of area-equivalent diameter coincide nicely for each time step  $t \in T$  and for each experiment except for IB3, where for increasing  $t$  the discrepancy between the medians of area-equivalent diameter computed from  $D_t$  and simulated data, respectively, tends to monotonously increase to approximately  $3.3 \mu\text{m}$ . The 95 percent quantiles of area-equivalent diameter computed from measured data in  $D_t$  are slightly larger than those computed from simulated data, especially for time steps  $t \in T$  in the middle of the experiments, see Figure 20.

## 5 Discussion

The present paper introduces a probabilistic modeling approach for hydrophobic agglomeration processes, where the bivariate probability densities  $f_{pp}^*$  and  $f_{agg}^*$  of the descriptor vectors (of area-equivalent diameter and aspect ratio) of primary particles and agglomerates, respectively, do not depend on the particular time step  $t \in T$ . For primary particles, this can be justified by the assumption that primary particles, i.e., single solid particles, do not change their sizes and shapes throughout the agglomeration experiments. On the other hand, the assumption that the joint density  $f_{agg}^*$  of area-equivalent diameter and aspect ratio of agglomerates does not change over time requires the presence of certain specific circumstances as, for example, the breaking of large agglomerates due to shear forces. This prevents agglomerates from growing unboundedly large, which in turn is necessary for an equilibrium in the size distribution of agglomerates.

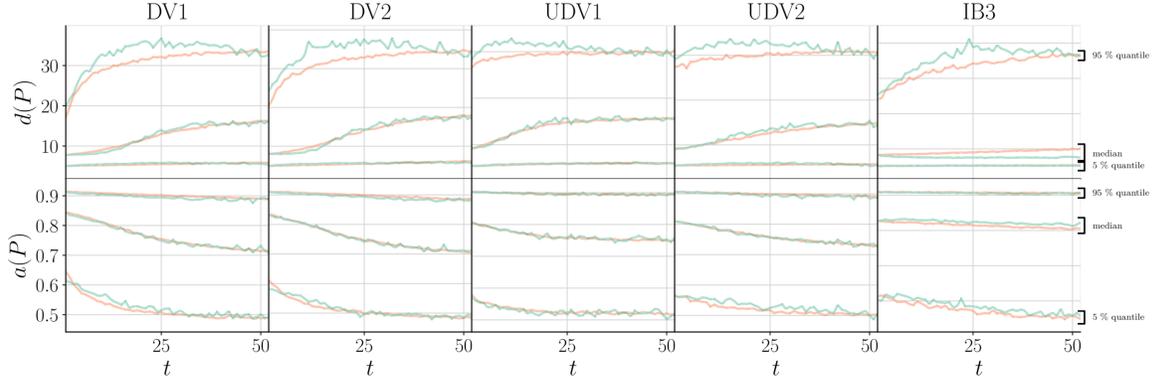


Figure 20: Median, as well as 5 and 95 percent quantiles of area-equivalent diameter (top) and aspect ratio (bottom) computed for  $D_t$  (green) and simulated data drawn from  $f_{m_t^*}^*$  (orange), respectively, for each time step  $t \in T$ .

The results obtained in this paper show that the fitted mixtures  $f_{m_t^*}^* = (1 - m_t^*)f_{pp}^* + m_t^*f_{agg}^*$  of the densities  $f_{pp}^*$  and  $f_{agg}^*$ , with appropriately chosen (time-dependent) mixing ratios  $m_t^*$ , manage to accurately describe the underlying data in  $D_t$  at each of the 52 time steps  $t \in T$ . Specifically, for the statistics considered in Section 4.5, i.e., for correlation coefficients, quantiles, mean values and standard deviations, only small discrepancies were observed between the respective quantities computed for data from  $D_t$  and simulated data drawn from  $f_{m_t^*}^*$ , for each time step  $t \in T$ , see Figures 18, 19 and 20. Besides this, our modeling approach satisfies some necessary consistencies across the experiments, e.g., the fitted probability densities  $f_{pp}^*$  being similar to each other, which is caused by the fact that all experiments use the same type of primary particles, see Figure 16, irrespective of the initial state of particle dispersion.

It should be emphasized that for the following two reasons the model is most likely not overfitted to measured data: on the one hand, the probability densities  $f_{pp}^*$  and  $f_{agg}^*$  are parametric and were only fitted by means of data from the first and a few time steps at the end of the agglomeration experiments. Consequently, they were not explicitly fitted for all time steps. Nevertheless, the evaluation results show good agreement across all time steps, indicating that the fits capture the distribution underlying the data quite well. A similar validation approach is also commonly used in neural networks to investigate overfitting, as these models are often trained iteratively and—due to their even larger number of parameters—tend to exhibit overfitting more strongly. On the other hand, the time-dependent mixing ratios  $m_t^*$  are modeled by a parametric function  $\gamma_{a,b,\lambda} : T \rightarrow [a, b]$  with merely three parameters which are closely related with the dynamics of the agglomeration experiments. Specifically, the parameters  $a$  and  $b$  represent the fractions of agglomerates at the beginning and the end of the experiment, respectively. The third parameter  $\lambda$  indicates whether the monotonous increase from  $a$  to  $b$  is linear over the course of the experiments, or whether the increase of the fraction of agglomerates is stronger at the beginning and weaker at the end of the experiment, or vice versa. Beyond their role in modeling the agglomeration process, the parameters  $a, b, \lambda$  can be utilized to guide process optimization. For instance, the low-parametric description of the agglomeration behavior by  $a, b, \lambda$  facilitates correlation with relevant process parameters, such as temperature, mixing speed or reactant concentrations, for example, by means of regression. This enables a systematic identification of dependencies between process parameters and the model parameters  $a, b, \lambda$ . Once these quantitative relationships are established, they can be leveraged for model-based process control and optimization, see, e.g., [30].

Recall that some agglomeration experiments considered in this paper dispersed the “feed-suspension” by ethanol and ultrasound, namely DV1, DV2 and IB3, whereas in experiments UDV1 and UDV2 no ultrasound was used. This leads to the assumption that the fitted fraction of agglomerates  $a^*$  at time step  $t = 1$ , i.e., 30-40 seconds after the “feed-suspension” is added to the tank, should be higher for UDV1 and UDV2 than for DV1, DV2 and IB3. This assumption is confirmed by the fact that the value obtained for  $a^*$  is approximately equal to 0.01 for experiments, where ultrasound was applied, and approximately equal to 0.11 for experiments, where no ultrasound was applied, see Table 7. Note that the smallest value for  $a^*$  is 0.0078 and has been obtained for experiment IB3.

Comparing the results obtained for the pairs of repeated experiments DV1/DV2 and UDV1/UDV2, it is expected that approximately the same fraction of agglomerates  $b^*$  is present at the ends of each pair of experiments at time step  $t = 52$ , as these experiments use the same energy dissipation rate associated with the Rushton turbine. One could

also expect that the value of  $b^*$  could be slightly higher for UDV1/UDV2, since  $a^*$  is larger for these experiments as mentioned above. Surprisingly, it is the other way around: the mean value of  $b^*$  is smaller for UDV1/UDV2, being equal to 0.64, than for DV1/DV2, where the mean value of  $b^*$  is equal to 0.81, see Table 7. Since the experiments DV1/DV2 and UDV1/UDV2 do not differ from each other except for the use or non-use of ultrasound, and because their fitted probability densities  $f_{pp}^*$  and  $f_{agg}^*$  are similar, one can argue that a more dispersed “feed-suspension” leads to a higher fraction of agglomerates at the end of the experiment.

This is likely a consequence of more agglomerates, i.e., large particles, being present at the beginning of experiments UDV1 and UDV2. Recall that large particles are more likely to immobilize at the gas-liquid interface, see Section 2.2.2. As a result, a higher number of primary particles, mostly part of an agglomerate, immobilize over the course of an experiment, where no ultrasound was used. Given that the same quantity of primary particles was used in each of the considered experiments in Table 1, this indicates a lower number of primary particles (including those in agglomerates) remaining under turbulence, i.e., not being immobilized, at the end of the experiments UDV1 and UDV2 compared to DV1 and DV2. Recall that agglomeration saturates when a certain minimum number/volume-based fraction of particles is reached, see Section 2.1. This fraction is consistent across the experiments DV1, DV2, UDV1 and UDV2 due to the uniform energy dissipation rate of 0.3 and the same volume of the suspension in the stirred tank, see Table 1. Since the probability densities  $f_{pp}^*$  and  $f_{agg}^*$  of size and shape descriptors for primary particles and agglomerates, respectively, are similar, the expected number of primary particles in an agglomerate is also similar across the experiments DV1, DV2, UDV1 and UDV2. Consequently, the higher number of primary particles (including those in agglomerates) remaining under turbulence at the end of the experiments DV1 and DV2 in comparison to the experiments UDV1 and UDV2, necessitates a higher fraction  $b^*$  of agglomerates at the respective end. This is required to reach the minimum number/volume-based fraction of particles for which agglomeration saturates.

Furthermore, one can compare the results obtained for experiment IB3 with those of DV1 and DV2, which differ in their energy dissipation rates that are associated with the two different stirrer types. Specifically, IB3 used an IB stirrer, whereas the experiments DV1 and DV2 both used a Rushton-turbine. As already mentioned above, the value of  $a^*$  is similar in each case due to the use of ultrasound. However, as we pointed out earlier  $a^*$  is slightly smaller for IB3 than for DV1 and DV2, as expected in Section 3.5.2, confirming that the higher energy dissipation rate in DV1 and DV2 results in a higher fraction of agglomerates within just 30-40 seconds after the “feed-suspension” is added to the tank. Moreover, for experiment IB3 the fitted value of  $b^*$  is equal to 0.25 and thus significantly smaller than the values of  $b^*$  obtained for experiments DV1 and DV2. However, the fitted value of  $\lambda^*$  for IB3 is equal to 0.3, which results in a function  $\gamma_{a^*,b^*,\lambda^*}$  of mixing ratios that does not seem to be saturating towards a value close to  $b^*$ , but a value that is larger than  $b^*$  by a non-negligible amount of (non-available) time steps  $t > 52$  as mentioned at the end of Section 4.4. This suggests extrapolating the mixing ratios  $\hat{m}_t$  for  $t > 52$  by means of those fitted from Eq. (33), or running an experiment under the same conditions for a longer time horizon, which could be realized in future research. By this, one could check if higher fractions of agglomerates are achievable when using an IB stirrer. For the remaining experiments DV1, DV2, UDV1 and UDV2 the fitted values of  $\lambda^*$  are smaller than 0.12, which in each case indicates a strong increase of the fraction of agglomerates at early time steps and a weak increase at later time steps, with saturation towards the end of the respective experiment.

Regarding the optimization of agglomeration processes, like those investigated in the present paper, not only the maximization of the fraction  $b^*$  of agglomerates is of interest, but also ways to design the process such that the resulting agglomerates possess desired size and shape characteristics. Figure 16 shows that the bivariate density  $f_{agg}^*$  of area-equivalent diameter and aspect ratio of agglomerates depends heavily on the energy dissipation rate generated by different stirrer types, but much less on the use or non-use of ultrasound. Specifically, flows induced by a Rushton-turbine or an IB stirrer lead to different bivariate densities  $f_{agg}^*$  of agglomerates, where two main distinctions can be identified. First, the mode of the area-equivalent diameter is at about 28  $\mu\text{m}$  for agglomerates resulting from the usage of an IB stirrer, whereas the modes of the area-equivalent diameter for agglomerates resulting from experiments employing a Rushton-turbine range from 16.5 to 18.5  $\mu\text{m}$ . Furthermore, the (negative) correlation of area-equivalent diameter and aspect ratio of agglomerates resulting from the usage of an IB stirrer is slightly stronger than for agglomerates resulting from using a Rushton-turbine, see Figure 12. However, the marginal distribution of aspect ratio shows minimal sensitivity to the choice of stirrer type and the associated energy dissipation rate, see the bottom row of Figure 16. On the other hand, the use of an IB stirrer, instead of a Rushton-turbine, results in larger agglomerates, but likely a lower fraction of agglomerates at the end of the experiment.

As indicated in the beginning of this section, our modeling approach can also be applied to other agglomeration processes, where particles can be partitioned into primary particles and agglomerates. However, note that for the

experiments considered in the present paper we assumed that no agglomerates are present in the set  $D_1$  of descriptor vectors obtained for the first time step  $t = 1$ . Therefore, it is important that the fraction of agglomerates at the beginning of an experiment that is intended to be modeled is not too large, i.e., particles are well dispersed initially. Table 7 shows that the fractions of agglomerates  $a^*$  at the beginning of experiments DV1, DV2, UDV1, UDV2 and IB3 are slightly larger than zero, which is contrary to our initial assumption that no agglomerates are present at time step  $t = 1$ . Thus, the small fractions  $a^*$  of agglomerates encompassed within the dataset  $D_1$  observed at time step  $t = 1$  are erroneously considered for fitting  $f_{pp}^{(0)}$ . Since we used  $f_{pp}^{(0)}$  to obtain the dataset  $D_{agg}$  given in Eq. (20), our assumption made regarding time step  $t = 1$  also affects the fit of  $f_{agg}^{(0)}$ . However, as we have shown, e.g., in Figure 14, the iterative adjustments explained in Section 3.4.2 can make up for a non-ideal initial fit  $f_{pp}^{(0)}$  created under the slightly perturbed circumstances mentioned above. Specifically, Figure 14 indicates that in the course of iterations subsets of  $D_{end}$  are found that are more representative for particle descriptor vectors of agglomerates than  $D_{agg}$ . This leads to better fits of the bivariate probability density  $f_{agg}$  of agglomerates which in turn leads to better fits of the corresponding density  $f_{pp}$  of primary particles. Nevertheless,  $f_{pp}^{(0)}$  should be somewhat close to the actual distribution of primary particles, i.e., the fraction of agglomerates at the beginning of an experiment must not be too large.

A key question is whether the methodology presented in this paper can be adapted to more complex particle systems beyond. A relevant extension could be to move from a monodisperse to a polydisperse particle feed, where the descriptors of primary particles follow multimodal distributions. In this case, the parametric model for the primary particles must be adjusted. Instead of a unimodal bivariate probability distribution, a mixture of multiple unimodal bivariate distributions may be required. If such a model can be established, all other steps of the algorithm, including the iterative procedure for distinguishing between primary particles and agglomerates, remain applicable without modification. This suggests that the presented methodology is adaptable to polydisperse particle feeds and potentially to other complex particle systems.

Building up on the methodology developed in this paper, future research may extend the current set of process parameters, namely the energy dissipation rate and the use or non-use of ultrasound, for investigating how various combinations of process parameters impact the effectiveness of hydrophobic agglomeration processes. This would entail establishing quantitative relationships that connect process parameters with characteristics of the resulting agglomerates, such as the mean value or variance of size and shape descriptors as well as the finally achieved fraction of agglomerates. By means of these relationships, the following inverse problem can be studied: for desired specifications of agglomerate characteristics the values of process parameters which are most likely to result in agglomerates exhibiting these predefined characteristics need to be determined. In this way, one can identify promising process parameters without having to carry out a large number of real laboratory experiments.

## 6 Conclusion

This study has introduced a suitable probabilistic modeling approach for hydrophobic agglomeration processes, where size and shape descriptors of primary particles and agglomerates follow fixed densities  $f_{pp}^*$  and  $f_{agg}^*$ , which do not depend on the particular stage of the agglomeration process. The stability of these densities throughout the agglomeration process confirms certain assumptions about process behavior, such as the breaking of large agglomerates due to shear forces, which prevents agglomerates from growing unboundedly large and thus contributes to an equilibrium in the size distribution of particles.

The obtained results demonstrate that the fitted mixtures  $f_{m_t^*}^* = (1 - m_t^*)f_{pp}^* + m_t^*f_{agg}^*$  of these densities, with suitably chosen time-dependent mixing ratios  $m_t^*$ , accurately capture the underlying data on area-equivalent diameter and aspect ratio of particles at each time step. While minor discrepancies were observed for certain statistics, overall, our model performed well in capturing the dynamics of agglomeration processes across various kinds of experiments.

Moreover, the discussion conducted in the previous section highlighted insights into the impact of process parameters, such as the use or non-use of ultrasound and the energy dissipation rate, on the agglomeration process. Specifically, experiments employing a more dispersed “feed-suspension” resulted in higher fractions of agglomerates at the end of the experiments. Furthermore, the energy dissipation rate associated with a specific stirrer type significantly affected the resulting agglomerate characteristics. Flows induced by an IB stirrer led to larger agglomerates with a slightly stronger negative correlation of area-equivalent diameter and aspect ratio compared to those obtained for a Rushton-turbine. On the other hand, the experiment which used an IB stirrer led to a lower fraction of agglomerates at the end of the experiment, confirming distinct agglomeration dynamics induced by the stirrer type with a

corresponding energy dissipation rate.

In view of these findings, future research could focus on extending the set of process parameters and establishing quantitative relationships between process parameters and characteristics of resulting agglomerates. By addressing the inverse problem of agglomeration process optimization, promising specifications of process parameters for achieving desired agglomerate characteristics can be identified without the need for extensive laboratory experiments. Thus, the present paper not only deepens our understanding of agglomeration dynamics, but also paves the way for practical advancements in industrial agglomeration processes, ultimately leading to cleaner, higher-quality products and more efficient suspension separation methods based on tailored agglomerates.

## Acknowledgement

The authors gratefully acknowledge funding by the German Research Foundation (DFG) within SPP 2364 under grants PE 1160/37-1 and SCHM 997/47-1 and for supporting the Collaborative Research Center CRC 920 (Project ID 169148856 – Subproject B04).

## References

- [1] S. Taniguchi, A. Kikuchi, T. Ise, N. Shoji, Model experiment on the coagulation of inclusion particles in liquid steel, *ISIJ International* 36 (1996) 117–120.
- [2] P. Knüpfer, L. Ditscherlein, U. A. Peuker, Nanobubble enhanced agglomeration of hydrophobic powders, *Colloids and Surfaces A: Physicochemical and Engineering Aspects* 530 (2017) 117–123.
- [3] F. Heuzeroth, J. Fritzsche, U. A. Peuker, Wetting and its influence on the filtration ability of ceramic foam filters, *Particuology* 18 (2015) 50–57.
- [4] J. Nicklas, U. A. Peuker, Agglomeration of fine hydrophobic particles: 1D and 2D characterization by dynamic image analysis of in-line probe data, *Powder Technology* 426 (2023) 118685.
- [5] F. Gruy, M. Cournil, P. Cugnet, Influence of nonwetting on the aggregation dynamics of micronic solid particles in a turbulent medium, *Journal of Colloid and Interface Science* 284 (2005) 548–559.
- [6] K. van Netten, D. J. Borrow, K. P. Galvin, Fast agglomeration of ultrafine hydrophobic particles using a high-internal-phase emulsion binder comprising permeable hydrophobic films, *Industrial & Engineering Chemistry Research* 56 (2017) 10658–10666.
- [7] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, Í. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, SciPy 1.0 Contributors, SciPy 1.0: Fundamental algorithms for scientific computing in Python, *Nature Methods* 17 (2020) 261–272.
- [8] R. Nelsen, *An Introduction to Copulas*, Springer, 2006.
- [9] F. Kruis, K. Kusters, The collision rate of particles in turbulent flow, *Chemical Engineering Communications* 158 (1997) 201–230.
- [10] X. Wang, X. Feng, C. Yang, Z.-S. Mao, Energy dissipation rates of Newtonian and non-Newtonian fluids in a stirred vessel, *Chemical Engineering & Technology* 37 (2014) 1575–1582.
- [11] C. Tomasi, R. Manduchi, Bilateral filtering for gray and color images, in: *Proceedings of the sixth International Conference on Computer Vision (IEEE Cat. No. 98CH36271)*, IEEE, 1998, pp. 839–846.
- [12] R. C. Gonzales, R. E. Woods, *Digital Image Processing*, Pearson Education, 2010.
- [13] O. Furat, T. Leißner, K. Bachmann, J. Gutzmer, U. A. Peuker, V. Schmidt, Stochastic modeling of multidimensional particle properties using parametric copulas, *Microscopy and Microanalysis* 25 (2019) 720–734.

- [14] S. Al-Thyabat, N. Miles, An improved estimation of size distribution from particle profile measurements, *Powder Technology* 166 (2006) 152–160.
- [15] L. Held, D. S. Bové, *Applied Statistical Inference: Likelihood and Bayes*, Springer, 2014.
- [16] O. Barndorff-Nielsen, Exponentially decreasing distributions for the logarithm of particle size, *Proceedings of the Royal Society of London. Series A* 353 (1977) 401–419.
- [17] M. Abramowitz, I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Vol. 55, US Government printing office, 1968.
- [18] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, J. Wiley & Sons, 2003.
- [19] D. W. Scott, *Multivariate Density Estimation: Theory, Practice, and Visualization*, J. Wiley & Sons, 2015.
- [20] C. Robert, G. Casella, *Monte Carlo Statistical Methods*, Springer, 2004.
- [21] R. M. Karp, An algorithm to solve the  $m \times n$  assignment problem in expected time  $O(mn \log n)$ , *Networks* 10 (1980) 143–152.
- [22] G. Loaiza-Ganem, J. P. Cunningham, The continuous Bernoulli: fixing a pervasive error in variational autoencoders, in: *Proceedings of the 33rd International Conference on Neural Information Processing Systems*, Curran Associates Inc., Red Hook, NY, USA, 2019, pp. 13287–13297.
- [23] G. S. Mudholkar, D. K. Srivastava, Exponentiated Weibull family for analyzing bathtub failure-rate data, *IEEE Transactions on Reliability* 42 (1993) 299–302.
- [24] P. R. Tadikamalla, A look at the Burr and related distributions, *International Statistical Review / Revue Internationale de Statistique* 48 (1980) 377–344.
- [25] N. L. Johnson, S. Kotz, N. Balakrishnan, *Continuous Univariate Distributions*, Vol. 2, J. Wiley & Sons, 1995.
- [26] R. D. Gupta, R. C. Gupta, Analyzing skewed data by power normal model, *TEST* 17 (2008) 197–210.
- [27] K. K. Talukdar, W. D. Lawing, Estimation of the parameters of the Rice distribution, *The Journal of the Acoustical Society of America* 89 (1991) 1193–1197.
- [28] K. Siekierski, Comparison and evaluation of three methods of estimation of the Johnson SB distribution, *Biometrical Journal* 34 (1992) 879–895.
- [29] S. Gündüz, M. Ç. Korkmaz, A new unit distribution based on the unbounded Johnson distribution rule: the unit Johnson SU distribution, *Pakistan Journal of Statistics and Operations Research* 16 (2020) 3.
- [30] O. Furat, M. Masuhr, F. E. Kruis, V. Schmidt, Stochastic modeling of classifying aerodynamic lenses for separation of airborne particles by material and size, *Advanced Powder Technology* 31 (6) (2020) 2215–2226.