ARTIFICIAL FILTER CAKE GENERATION: DIGITAL TWINS VIA STOCHASTIC 3D MODELING BASED ON μ-CT IMAGE DATA

PHILLIP GRÄFENSTEINER[1], ERIK LÖWER[2], ORKUN FURAT[1],
URS A. PEUKER[2], VOLKER SCHMIDT[1]

[1]*Institute of Stochastics, Ulm University, 89069 Ulm, Germany*

[2]*Institute of Mechanical Process Engineering and Mineral Processing,
Technische Universität Bergakademie Freiberg, 09599 Freiberg, Germany*

ABSTRACT. In order to link the properties of a feed stream to those of a product in separation operations such as cake filtration, a comprehensive database is required which often cannot be achieved through laboratory measurements alone. For this reason, a stochastic 3D model for the generation of virtual filter cake structures is developed and calibrated to tomographic image data of experimentally built filter cake structures. In this way, digital twins of real particles can be simulated using mixed Gaussian random fields on the unit sphere in $\mathbb{R}^3$, which are then spatially arranged to form a three-dimensional artificial filter cake. Its 3D morphology is validated with respect to geometric descriptors that were not used for model fitting, such as tortuosity, constrictivity and specific surface area of pore space. In future work, using the stochastic 3D model developed in the present paper, a large database of systematically varied artificial filter cakes will be generated by adjusting interpretable model parameters, allowing for transfer functions to be determined and structure-property relationships to be statistically investigated.

1. INTRODUCTION

In process engineering, it is often necessary to separate liquids from particulate solids [1–3], for which one option is cake filtration. In this process, a cake of particles is formed on a support fabric by which further particles are separated from the liquid. This filter cake continues to grow as the process time increases, until the increasing pressure drop across the bed, requiring the entire filter cake to be removed. A new separating cake structure is then rebuilt in a next process cycle. Cake filtration is based on the micro-process of pore flow, which is determined by the permeability or its reciprocal, the specific filter cake resistance [4, 5]. Often the permeability is linked to just a single geometric particle descriptor, namely the Sauter diameter [6] by the Carman-Kozeny equation [7] which shows, in many cases, a high numerical error [8]. However, the entire 3D morphology of the particle collective participating in the process influences associated effective properties such as capillary pressure, permeability, residual moisture and saturation. To design a solid-liquid separation process, these properties need to be estimated, which is done using various empirical, semi-empirical or analytical models [9–11]. In these process models, information on particle geometry is often reduced to a single aggregated descriptor, which provides only limited quantitative information about the particle system and the corresponding structure-property relationships. Moreover, these aggregated descriptors alone are insufficient to fully describe or model process-related effects in detail or in some cases provide contradictory dependencies that cannot be resolved even in parametric fits [12–14]. It is important to consider multi-dimensional information on particle geometry to accurately model

and understand the particle system as well as the corresponding process [15]. The interaction of various descriptors of the micro-structure of a filter cake is decisive for the performance of the process, but this cannot be aggregated into a single value. In particular, the internal structure of the cakes cannot be described by single-value descriptors but rather by distributions of descriptors [14, 16].

Previous modeling attempts [17–23] have been proposed to capture the relationship between relevant geometric descriptors and resulting process properties such as particle size and permeability. This results in additional model parameters such as modified porosity functions. However, it is often difficult or outright impossible to isolate the influence of a single particle shape or size descriptor without altering other relevant descriptors of the particle system. For example, since the porosity of a filter cake varies when the individual particle size distribution is changed, it is not possible to say with certainty whether an observed effect is due to the change in porosity or the change in particle size and/or shape distribution. In order to isolate the influence of porosity, it would be necessary to vary the porosity over a wide range independently of other influences. In [24], structure-property relationships were analyzed experimentally, where it has been concluded that only the porosity and information on the particle size and shape distributions are geometric variables influencing the effective properties of a filter cake. However, this theory is difficult to test experimentally, as porosity and packing structure are varied simultaneously with, for example, particle shape.

In our research, we therefore analyze shape and size descriptors of individual particles in detail, utilizing tomographic image data of filter cakes which consist of spherical glass beads and of quartz particles obtained from 3D measurements by X-ray computed tomography (μ-CT), see Figure 1. The filtration tests correspond to conventional laboratory tests in a pressurised nutsch with adaptations to the measuring principle of microscopy. From this comprehensive analysis, we obtain particle-discrete segmented datasets which are used to establish a descriptor vector for each particle.
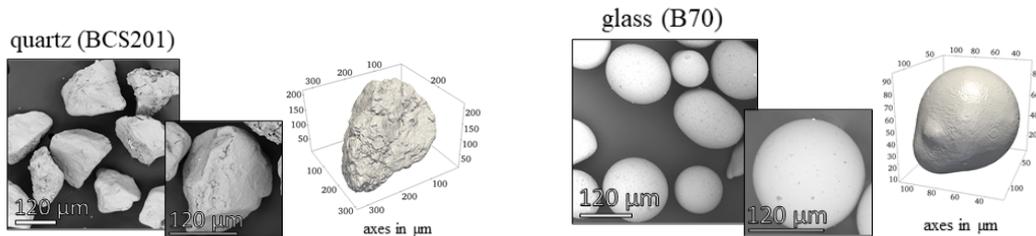
FIGURE 1. Scanning electron microscope images of quartz (left) and glas (right) particles, with 3D renderings of individual particles from μ-CT measurements.

Unfortunately, the variety of available tomographic image data is not large enough to reliably determine transfer functions [25] for the relationship between geometric particle descriptors and filter cake properties we measured in our experiments. One approach to solving this problem is to expand the database by model-based generation of virtual but realistic filter cake structures. In this way, transfer functions can be determined to infer filter cake properties and ultimately process properties from geometric particle descriptors.

Thus, based on tomographic image data, we develop a spatial stochastic model [26, 27] that is able to generate virtual 3D morphologies of filter cake structures which are statistically similar to real filter cake structures in an appropriate sense. More precisely, the artificial structures are validated by their real-life counterparts in terms of geometric descriptors that were not used for model fitting, such as constrictivity, porosity, tortuosity and specific surface area. Our model operates in two steps. At first, virtual individual particles are generated by means of a mixed

Gaussian random field on the unit sphere in $\mathbb{R}^3$ [28], which is calibrated to the particles observed in the segmented tomographic image data using spherical harmonics functions. Such a modeling approach has previously been used in [25, 29] to model the shape of particles in cathodes of lithium-ion batteries. In the present paper, we extend this methodology by deploying a rejection sampling scheme [30, 31] based on the joint distribution of particle volume and sphericity, which is modeled by means of parametric copula functions [30].

This approach has two main advantages. First, the rejection sampling scheme allows us to reliably generate particles of a given bivariate size and shape distribution, even when particles exhibit a pronounced aspherical shape. The latter is often difficult to reproduce with models based on spherical harmonics. On the other hand, parametric modeling by means of a bivariate copula function allows us to vary the desired distribution of size and shape as well as their interdependence systematically by adjusting the corresponding parameters. In this way it is possible to generate ensembles of particles that each follow a different size and shape distribution from the same model type without the need for new tomographic image data. In a second step, these individual particles are then packed into artificial filter cake structures using a forced-bias algorithm [32, 33].

By means of the presented model we lay the foundation for a large-scale (data-driven) investigation of structure-property relationships for filter cake structures. The model is able to generate virtual filter cake structures whose individual particles follow a wide range of shape and size distributions, so that a large database of varying artificial filter cake structures can be generated and analyzed with respect to properties of their pore space. In future work, such an extensive database will be used to quantify the relationships between descriptors of individual particle geometry and effective properties of the resulting filter cakes such as permeability or de-watering behavior.

The rest of this paper is organized as follows. In Section 2, the material samples and imaging techniques considered in this paper are described. Section 3 deals with the segmentation of gray-scale images, whereas Sections 4 and 5 introduce the single particle model and the filter cake model, respectively. The results obtained in this paper are presented and discussed in Section 6. Finally, Section 7 concludes.

2. Materials and Imaging

In order to create a limited but sufficient database for modeling, we selected (i) simple but application-oriented material types that have certain process properties during filtration and are therefore suitable for (ii) laboratory experiments on cake filtration and also (iii) allow for three-dimensional, non-destructive measurements of almost the entire cake structure with a resolution limit below the particle size to be able to quantify the particle descriptor vectors considered in this paper. In particular, in Section 2.1 we discuss the selection of materials as well as their characteristic properties and describe the filtration process, whereas in Section 2.2 we explain the image measurement principle and how we adapted the process to it.

2.1. **Materials and filtration experiments.** In tomographic studies, the smallest adequate sample sizes are required to avoid extreme exposure times and high acceleration voltages, while the required resolution can only be achieved with smaller sample sizes due to the limited geometric magnification. For this reason we developed a downscaled pressure nutsch, which has been reduced in size compared to conventional laboratory equipment. The laboratory test rig is a down-scaled pressure nutsch for cake forming filtration according to the German guideline VDI 2762-2 [34]. It was reduced by $1/100$ in filter area which results in a cake formation unit with a diameter of $5\,\mathrm{mm}$. The downscale procedure is described in [16] in more detail.

The two solid particle systems involved in our studies for filtration tests differ from each other in terms of particle size and shape as well as with respect to their material properties. Due to the wide distribution of particle properties, the resulting process characteristics in filtration

and cake structure vary accordingly. One solid used in the experiments as disperse phase is soda-lime glass consisting of spherical particles with size x between 70 and 110 µm (B70) (Sigmund Lindner GmbH, Germany), which is manufactured via dry spraying. The soda lime glass beads consist of 72.30 m.-% SiO_2, 13.30 m.-% Na_2O, 8.90 m.-% CaO and 4.00 m.-% MgO with other additional trace elements. With a density of 2.55 g/cm^3, the material has an X-ray mass attenuation coefficient of 1.49 cm^2/g at 20 keV [35]. The second solid, quartz sand with the commercial designation BCS201, was purchased from Strobel Quartzsand GmbH, Germany, in the fraction $x < 200$ µm and is usually used as filling material for construction. The main component of the second particle system is 99.1 m.-% SiO_2 with traces of Fe_2O_3, Al_2O_3 and TiO_2 with < 0.3 m.-% each. This results in an X-ray attenuation coefficient of 2.44 cm^2/g at 20 keV [35] with a density of 2.65 g/cm^3. All densities of the solids were determined using gas pycnometry (Micromeritics Multivolume Pycnometer AccuPyc II 1340).

All experiments begin with the preparation of the suspension. The continuous phase of the suspension is an aqueous solution of de-ionized water, potassium iodide (purity > 99.9 m.-%, Carl Roth, Germany) and glycerol ($C_3H_8O_3$, purity > 98.5 m.-%, Carl Roth, Germany). This composition with a density of 1.10 g/cm^2 results in an attenuation coefficient of 1.68 cm^2/g at 20 keV [35]. The characteristic values for filtration, such as the masses of solids used per individual test m_S, the volume fractions of solids in the suspension φ, and the filtration pressure differences Δp applied, are listed in the appendix, see Table A1.

The suspension is stirred at 250 min^{-1} for 4 min while degassed at $p_U = 0.2$ bar. After filling the suspension into the nutsch and starting the vacuum pump, the filtrate is collected inside a condensate separator. The filter medium consists of a calendered polypropylene woven fabric (05-1010-SK 006, Sefar, Switzerland). For all measurements, filtration ends after gas breakthrough. In these tests, the flow cell is thoroughly weighed immediately after filtration and drying (24 h at 80 °C), so that the residual moisture RM can be determined. Further cake properties, like porosity ε, filter media resistance R_M, and filter cake resistance r_C can be derived from the recorded filtrate mass curves as a function of time, measured cake heights and cake masses according VDI 2762-2 [34], see Table A1.

2.2. **Image measurement principle.** Figure 2 schematically shows the measurement principle of the Zeiss Xradia 510 Versa microscope, consisting of a basic cone beam imaging µ-CT system and an extended X-ray microscope (XRM) used for subsequent measurements. With so-called µ-CT devices, the magnification is strictly geometric, since an optical system does not refract X-rays. For an X-ray microscope (XRM), after the scintillator, a magnifying optical unit transmits the visible light by means of different lenses that can be chosen individually. This design allows optical magnification in addition to the geometric magnification.
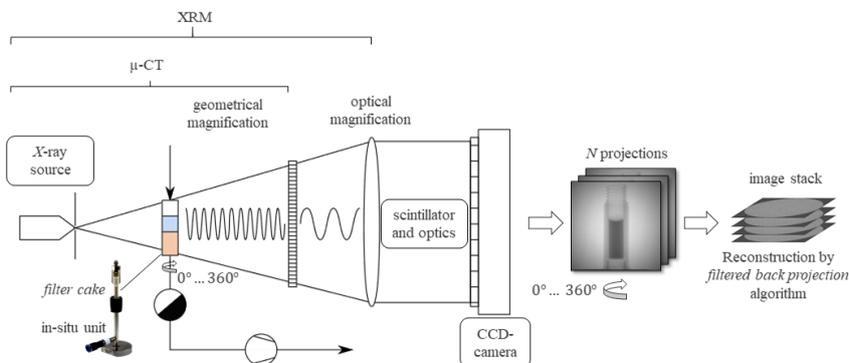
FIGURE 2. Scheme of the measurement principle deployed by the Zeiss Xradia 510 Versa microscope used for all experiments, where a conventional μ-CT device is followed by an X-ray microscope (XRM).

The sample, our filter cake structure within the in-situ filter cake unit, is placed between the X-ray source and the detector, where it is rotated by 360° during the measurement. Individual X-ray images (projections) of the sample are taken at certain angular distances so that after the measurement a dataset of N projections is obtained covering the entire range $0, \ldots, 360°$.

Using a mathematical reconstruction method, a three-dimensional volumetric image of the sample can be generated from this image information. In our case, this tomographic reconstruction is performed by means of the filtered back projection (FBP) algorithm [36], which is implemented in the ZEISS Xradia reconstruction software (Xradia XMReconstructor, v14). This includes an automatic centre shift and a beam hardening correction with a factor of 0.05, where the raw data is smoothed by a Gaussian filter with $\sigma = 0.7$. The measurement parameters of both filter cakes mentioned in Section 2.1 can be found in the appendix, see Table A2.

As a result of a measurement, the sample is represented by a three-dimensional array of cubic voxels, after reconstruction of the projection images at different angles. In particular, after reconstruction of the projection images, Figures 3a and 3d show 2D slices of the filter cakes built by glass beads and quartz sand, respectively.
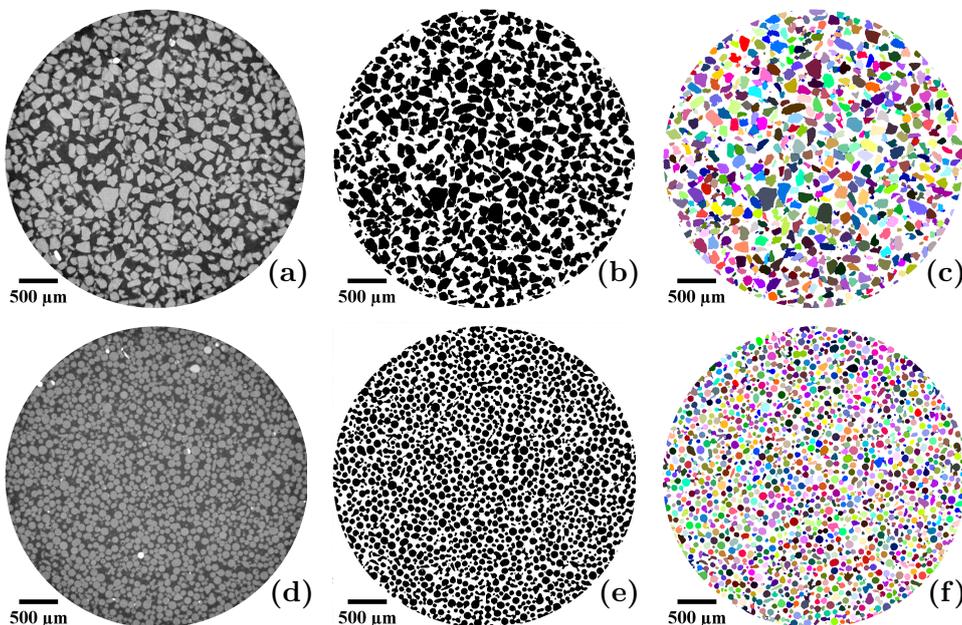
FIGURE 3. Left (a, d): Raw tomographic image data. Middle (b, e): Binarized image data. Right (c, f): Particle-discrete segmentation obtained from adapted watershed algorithm. Top (a–c): quartz. Bottom (d–f): glass

3. Image Segmentation

3.1. **Binarization.** The raw image data contains gray-scale information which makes it difficult to distinguish whether voxels belong to the pore space or to a particle. We therefore binarize the image data so that foreground voxels correspond to particles and background voxels to pore space. Formally, we consider a binary 3D image to be a mapping from the infinite domain $\mathbb{Z}^3 = \{\ldots, -1, 0, 1, \ldots\}^3$ to $\{0, 1\}$, where the value 1 corresponds to the foreground, and the value 0 to the background. While digital images are only defined over a finite subset of $\mathbb{Z}^3$, they can be extended to $\mathbb{Z}^3$ in an appropriate sense to fit into this framework, usually by assigning a constant value to voxels outside the finite domain. The binarization of gray-scale images is achieved by means of the algorithm `Intermodes` [37] implemented in the image processing software Fiji [38]. This algorithm assumes that the histogram of the underlying gray-scale values

within the image data is bimodal, which is the case for the tomographic image data considered in this paper. If the values of the two local maxima are given by $s_1, s_2 \in \mathbb{R} = (-\infty, \infty)$, then the threshold for binarization is set to $t = (s_1 + s_2)/2$. Iterative smoothing of the histogram is used to determine these maxima. The resulting binarizations are visualized in Figures 3b and 3e.

3.2. **Particle-discrete segmentation.** With the binarized tomographic image data at hand, we now describe how to obtain a particle-discrete segmentation. More precisely, this type of segmentation is a mapping $S\colon \mathbb{Z}^3 \to \{0, 1, 2, \ldots\}$ that assigns a positive integer to every foreground voxel of the tomographic image data and 0 to voxels belonging to the background. Voxels that are assigned the same positive integer are considered to belong to the same particle. The procedure that we use for determining such a particle-discrete segmentation is based on the watershed algorithm [39]. For this, we consider the Euclidean distance transform of binary image data. More precisely, given a binary image $B\colon \mathbb{Z}^3 \to \{0, 1\}$, the Euclidean distance transform $D\colon \mathbb{Z}^3 \to [0, \infty)$ of B assigns every voxel $v \in \mathbb{Z}^3$ a value that corresponds to the shortest distance from v to the nearest background voxel, *i.e.*, the nearest voxel with value 0 in the binary image B. Formally, the Euclidean distance transform D is given by

$$D(v) = \inf\{\|w - v\| \colon w \in \mathbb{Z}^3, B(w) = 0\}, \tag{1}$$

for any $v \in \mathbb{Z}^3$, where $\|\cdot\|$ denotes the Euclidean norm on $\mathbb{R}^3$. Clearly, if $B(v) = 0$, then $D(v) = 0$ for any $v \in \mathbb{Z}^3$. Intuitively speaking, the watershed algorithm then treats D as a topological map, where water sources are placed at regional minima, and the map is flooded with water rising from these sources. The ridges, along which the water from distinct sources meet, induce a particle-discrete segmentation of the original binary image. The final number of regions detected by this algorithm is identical to the number of initially placed water sources, and thus to the number of regional minima of the Euclidean distance transform. Therefore, the watershed algorithm is often prone to oversegmentation and requires some adjustments. In particular, we make use of morphological reconstruction [40] to reduce the effect of oversegmentation [41], which has been used in [42] to get an appropriate particle-discrete segmentation of gray scale image data. For a gray-scale image $G\colon \mathbb{Z}^3 \to \mathbb{R}$, morphological reconstruction involves iterative morphological dilation $G \oplus E$ of G with a structuring element $E \subset \mathbb{Z}^3$, defined as

$$\big(G \oplus E\big)(v) = \max_{w \in E} G(v + w) \tag{2}$$

for any $v \in \mathbb{Z}^3$. After every iteration, the point-wise minimum of the dilated image and a predefined mask image $M\colon \mathbb{Z}^3 \to \mathbb{R}$ is taken. Note that in order to make sense of Eq. (2) for real world applications, the digital image, leading to G, needs to be extended to $\mathbb{Z}^3$ by its minimum value (instead of simply extending it with the value 0). Formally, we then set $G_0 = G$ and $G_k = \min\{M, G_{k-1} \oplus E\}$, repeating this procedure until $G_k = G_{k-1}$ for some integer $k \geq 1$. The resulting image G_k is the morphological reconstruction of the so-called marker image G using the structuring element E and the mask M. Various image processing tasks can be performed by use of morphological reconstruction with different choices of G and M, such as filling holes, detecting connected components, or finding regional minima or maxima [39]. For our application of reducing oversegmentation in the watershed algorithm, the marker image is chosen as $(1 - \alpha)D$, where D is the Euclidean distance transform defined in Eq. (1), and α is a manually chosen parameter. In our case, both the values of $\alpha = 0.25$ for quartz sand and $\alpha = 0.3$ for glass have been determined manually to yield good results. The mask image M is then set to be the Euclidean distance transform D. The structuring element E is a cube with a side length of 3 voxels. Intuitively speaking, this choice of marker and mask image causes small ridges of the topological map D to disappear, so that only significant ridges are considered, which counteracts oversegmentation. The watershed algorithm is then applied to the resulting morphological reconstruction, yielding a particle-discrete segmentation S, visualized in Figures 3c and 3f.

3.3. **Validation via laser diffraction measurements.** From the segmented image $S : \mathbb{Z}^3 \to \{0, 1, 2, \ldots\}$ we readily obtain a particle size distribution by counting the number of voxels that belong to each particle. Formally, this corresponds to counting the number of elements in each of the pre-images $S^{-1}(\{i\})$, $i \in \mathbb{N} = \{1, 2, \ldots\}$. The quality of the segmentation S is assessed by comparing this distribution to that obtained via laser diffraction measurements for both quartz and glass particles.

For this, the particles are dispersed in a 1 g/L aqueous TNPP solution ($C_{45}H_{69}O_3P$) and ultrasonicated for 2 min. Three laser diffraction measurements were taken and the resulting distributions (Fraunhofer approximation) were determined using Sympatec HELOS/Qixel-R5, which has a measuring range of 4.5 to 875 µm. Note that the particle size determined in this way is the diameter of the diffraction equivalent sphere.

Figure 4 shows the distributions of particle sizes, expressed by means of their volume-equivalent diameters, which have been acquired via image segmentation and laser diffraction measurements, for both quartz and glass particles. We observe that there is a high agreement between the distributions obtained by image segmentation and laser diffraction measurement for glass particles, and a slightly larger discrepancy between these distributions for quartz particles. It is worth noting that particle size distributions determined by laser diffraction measurements assume particles to be perfectly spherical, which can be part of the reason for the deviations observed in Figure 4 for the more aspherical quartz particles.
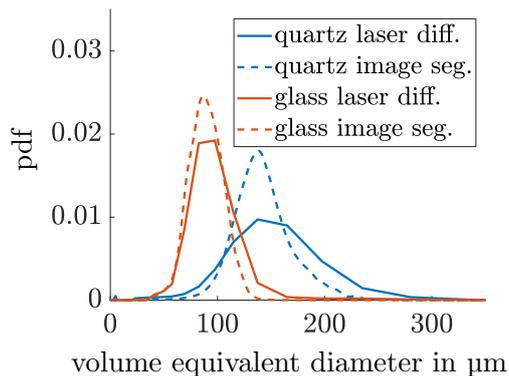
FIGURE 4. Probability density functions (pdf) of volume-equivalent diameters determined by laser diffraction (solid lines) and image segmentation (dashed lines).

4. SINGLE PARTICLE MODEL

In this section we describe the methods used to derive a stochastic model for the outer shape of individual quartz and glass particles. Our model is based on a spherical harmonics representation for generating random particle shapes as described in Section 4.1 below. This parametric model is calibrated separately to the tomographic image data of quartz and of glass particles, obtaining two distinct sets of parameters that correspond to the very different shapes of quartz and glass particles. The goodness of model fit is evaluated in Section 4.2. Afterwards, the two model instances are combined into a single model, where the size and shape of generated particles is controlled by a rejection sampling scheme. The advantage of this approach is that we can simulate particles of a desired size and sphericity distribution without the need to recalibrate the parameters of the spherical harmonics expansion. Instead, we can virtually vary the shapes of particles solely through the rejection sampling. This is greatly beneficial for subsequent (virtual) structural scenario analyses, where no tomographic image data is available for model

calibration, while varying the target distribution in rejection sampling is feasible without new tomographic measurements of different particle systems.

More precisely, the two initial model instances are combined by drawing on average half of the realizations from the model instance calibrated to quartz particles, and on average the other half from the model instance calibrated to glass particles. For this, the joint distribution of volume and sphericity for both tomographic image data and model realizations of quartz and glass particles is represented using a parametric copula fit, which is introduced in Section 4.3.1. The usage of copula functions allows us to model the dependency between particle sizes and shapes accurately while maintaining a high flexibility for the choice of parametric families for the univariate marginal distributions. Based on these fits, the joint probability density of volume and sphericity for the mixture of the two model instances is given as a linear combination of the individual joint probability densities. This linear combination is then used as the source distribution of the rejection-acceptance method, which allows us to reject or accept the drawn samples based on the desired joint distribution of particle volume and sphericity, as described in Section 4.3.2.

4.1. **Spherical harmonics representation.** The stochastic 3D model for single particles of both quartz sand and glass particles is based on the theory of spherical harmonics, following the approach developed in [29]. For this, we assume that the individual particles are star-shaped, which means that every particle has a so-called star point, from which the line segment to any other point of the particle is fully contained within the particle. Formally, we require that for every particle $P \subset \mathbb{R}^3$, there exists a point $s \in P$, such that for any other point $t \in P$, and any $\lambda \in [0,1]$, it holds that $s + \lambda(t-s) \in P$. The outer shell of such a particle can then be described by a radius function $r\colon S^2 \to [0, \infty)$, where $S^2 = \{u \in \mathbb{R}^3\colon \|u\| = 1\}$ is the unit sphere. For any direction $u \in S^2$, the value $r(u)$ is the distance from the star point to the particle boundary in direction u. In order to stochastically model the outer shell of the individual particles, it suffices to determine random functions from S^2 to $[0, \infty)$ which behave statistically similar to the radius functions of observed particles. For this, we use a suitably chosen family of random fields $X = \{X(u)\colon u \in S^2\}$ on the unit sphere S^2 [43]. In particular, we assume that X is the mixture of motion-invariant, *i.e.*, stationary and isotropic, Gaussian random fields [28]. The assumption of stationarity and isotropy leads to great reductions in model complexity, which simplifies the estimation of model parameters. On the other hand, by considering mixtures of Gaussian random fields we ensure flexibility that allows us to accurately represent a wide spectrum of particle shapes through the model, without adding significant complexity to model calibration. Formally, for any fixed integer $k > 1$, a mixture of Gaussian random fields is given by k (stochastically independent) Gaussian random fields $X_1, \ldots, X_k$ on S^2 and probabilities $p_1, \ldots, p_k \in [0,1]$ such that $\sum_{i=1}^k p_i = 1$. Then,

$$X = \begin{cases} X_1, & \text{with probability } p_1, \\ \vdots \\ X_k, & \text{with probability } p_k. \end{cases} \tag{3}$$

The number of components k, the mixing probabilities $p_1, \ldots, p_k$ as well as the parameters of each of the k Gaussian random fields $X_1, \ldots, X_k$ have to be estimated from image data. For more information on Gaussian random fields themselves and the estimation of parameters for mixtures of Gaussian random fields, we refer to [29]. In practice, any realization of the random field X needs to be discretized. This means that in order to generate a realization of X we first randomly choose one of the Gaussian random fields $X_1, \ldots, X_k$ according to the probabilities $p_1, \ldots, p_k$ and then evaluate the chosen X_i at a certain number $m > 1$ of grid points $u_1, \ldots, u_m \in S^2$, which corresponds to generating a realization of the m-dimensional (normally distributed) random vector $\big(X_i(u_1), \ldots, X_i(u_m)\big)$.

The model is calibrated separately to the datasets of quartz and of glass particles, yielding two distinct instances of the model with different parameter values, see Figure 5. In the following, we will denote the mixtures of Gaussian random fields calibrated to tomographic image data of quartz and glass particles by X_Q and X_G, respectively.

4.2. **Evaluating the fitted mixtures of Gaussian random fields.** In order to validate the fitted mixtures of Gaussian random fields X_Q and X_G, we consider the volume and the sphericity of each particle. Note that the sphericity is a quantity that measures the deviation of particle shape from that of a sphere. It is defined as the ratio between the surface area of a sphere with the same volume as the considered particle and the surface area of the particle. Since the sphere has the lowest ratio of surface area to volume, the sphericity always takes values between 0 and 1, where it is equal to 1 if the particle is a sphere and smaller than 1 for particles with non-spherical shapes.

Thus, both the volume and the surface area are needed in order to compute the sphericity of measured and simulated particles. By evaluating the mixed Gaussian random fields X_Q and X_G at gridpoints $u_1 \ldots, u_m \in S^2$, we obtain a triangular mesh of the unit sphere, where each triangle is defined by some triplet $(u_{i_1}, u_{i_2}, u_{i_3})$. Summation of the areas of these triangles yields an estimator for the surface area of particles. An estimator for the volume of particles is obtained by summing up the volumes of the tetrahedra defined by the quadruple $(u_{i_1}, u_{i_2}, u_{i_3}, o)$, where $o = (0, 0, 0) \in \mathbb{R}^3$ denotes the origin. In Figures 5a and 5b univariate probability densities of (number weighted) particle volume and sphericity are visualized for model realizations as well as for particles observed in tomographic image data.
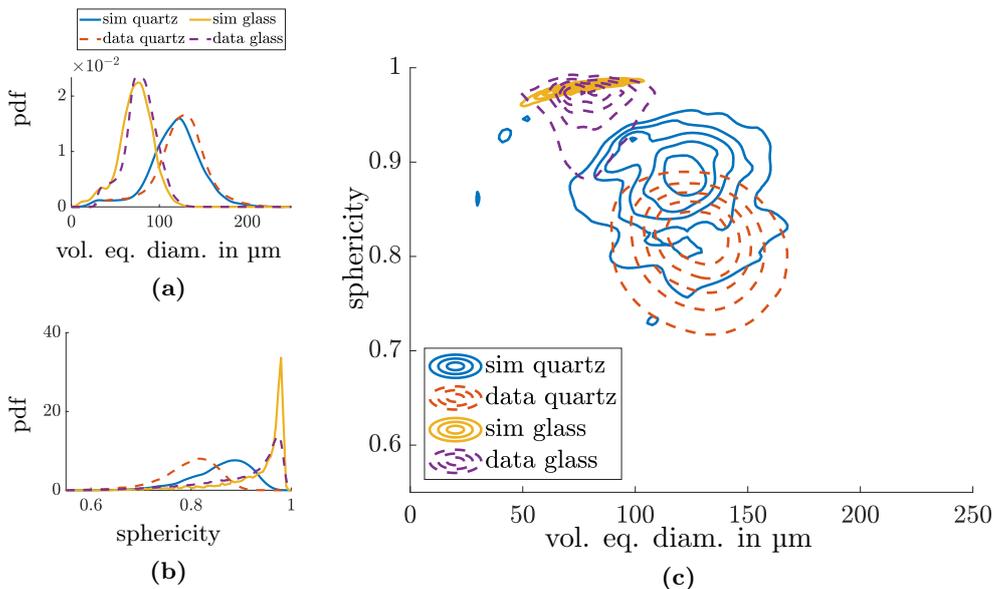
FIGURE 5. Number-weighted distributions of volume equivalent diameter and sphericity of measured and simulated particles of both glass and quartz particle systems: the univariate probability densities of (a) volume-equivalent diameter and (b) sphericity, and (c) their joint bivariate probability densities. For tomographic image data, 155890 glass and 57977 quartz particles have been considered. For simulated data, 50000 particles were generated in each case.

It turned out that the distribution of particle volumes observed in tomographic image data is nicely represented by the model for both the quartz and glass data sets, see Figures 5a. However, while the sphericity of the largely spherical glass particles is also well reproduced,

the sphericity of the aspherical quartz particles is less accurately captured by the model, see Figure 5b. Although the overall shapes of the density functions are similar, the distribution of the sphericity of simulated quartz particles is shifted towards higher values. We believe that a reason for this discrepancy might be the fact that particle shapes with rough edges are typically difficult to mimic accurately by means of the spherical harmonics approach. Moreover, Figure 5c shows that the dependency between the size and shape descriptors is not correctly reproduced by the model, in both cases. In the following, we therefore combine the model presented in Section 4.1 with a rejection sampling scheme.

4.3. **Copula-based rejection sampling.** As explained in Section 4.2, the mixtures of Gaussian random fields introduced in Section 4.1 are not able to represent the shape of the aspherical quartz particles sufficiently well. We therefore combine this modeling approach with a copula-based rejection sampling scheme [30, 31]. In the case of quartz particles, this will allow us to discard model realizations of particles which are too spherical, and only accept those realizations which exhibit a realistic sphericity value.

More precisely, we can readily sample particles whose volume and sphericity follow the bivariate probability density visualized in Figure 5c by solid blue contour-lines, where the samples are drawn from the random field X_Q, as explained in Section 4.1. However, we would like the volume and sphericity of our samples to follow the bivariate probability density of particle descriptors observed in tomographic image data, visualized by the dashed orange lines in Figure 5c. This can be achieved by repeatedly drawing samples from the random field X_Q (the volume and sphericity of which follow the bivariate density depicted in blue) and rejecting or accepting a sample with an appropriate probability, depending on the particular sphericity and volume of that sample. In order to determine the correct probability of rejection or acceptance, it is convenient to have an analytical representation of all involved probability densities. Therefore, we first fit a parametric function to the bivariate probability density of sphericity and volume for all cases, *i.e.*, for samples drawn from the models X_Q and X_G as well as for the tomographic image data of both quartz and glass particle systems. When modeling such multivariate distributions, a copula-based approach allows us to model the univariate marginal distributions and the correlation structure of the underlying one-dimensional particle descriptors separately. In this way, we are not restricted to the limited choice of conventional parametric families of bivariate probability densities.

Moreover, this approach allows us to combine the models X_Q and X_G into a single model for the individual particle shapes, which is able to produce particles that meet a wide range of predefined joint distributions of volume and sphericity. In this way, it is not necessary to recalibrate the model to tomographic image data of newly measured particles. Instead, one can generate virtual particles and, subsequently, virtual filter cake structures by merely defining the desired joint distribution of volume and sphericity. To this end, we will sample from the models X_Q and X_G with a probability of 0.5 each. That is, we consider the combined model $X_{Q,G}$ which is given by

$$X_{Q,G} = \begin{cases} X_Q, & \text{with probability } \frac{1}{2}, \\ X_G, & \text{with probability } \frac{1}{2}. \end{cases} \tag{4}$$

Note that the bivariate probability density of sphericity and volume for samples drawn from $X_{Q,G}$ is a linear combination of the corresponding bivariate probability densities associated with X_Q and X_G, respectively. Based on this mixed sampling approach, we employ rejection sampling in order to reproduce the desired joint distribution of sphericity and volume, for both the quartz and glass datasets.

Therefore, we first present some basic background of copula theory in Section 4.3.1, including how copulas can be utilized to derive and fit multivariate parametric probability distributions to data. Then, in Section 4.3.2, we show how the obtained parametric fits can be deployed for rejection sampling.

4.3.1. *Copula fits.* The parametric modeling approach that we deploy in this paper is based on the theory of copulas [30]. In this way, any probability density of a multi-dimensional random vector can be decomposed into a product of its univariate marginal probability densities and an appropriately chosen copula density, which captures information on the dependencies between the underlying one-dimensional random variables.

In the following, we will explain the copula approach specifically for the two-dimensional case, although it generalizes in a straightforward way to arbitrary dimensions. The idea is that instead of modeling bivariate densities directly, the fitting procedure is split into two steps. In the first step, the univariate densities are modeled by a parametric fit. The correlation structure of the underlying one-dimensional particle descriptors is then represented by a copula, which is modeled in a second fitting step by comparing the goodness-of-fit between a wide range of parametric copula families [44]. Finally, the univariate distributions and the copula are combined into a parametric model of the bivariate probability density. Such a modeling approach allows for great flexibility in modeling various types of dependency structures without restricting the modeling of the univariate distributions.

Formally, a bivariate copula $C: [0,1]^2 \to [0,1]$ is the (joint) cumulative distribution function of a random vector $U = (U_1, U_2)$ such that the marginal distributions of its components U_1 and U_2, are uniform on the unit interval $[0,1]$. Sklar's representation theorem (*cf.* Theorem 1.1 in [30]) states that for any bivariate distribution function $F: \mathbb{R}^2 \to [0,1]$ of a 2-dimensional random vector there exists a copula C such that F can be written in terms of its marginal distribution functions F_i, $i = 1, 2$, and C as

$$F(x_1, x_2) = C\big(F_1(x_1), F_2(x_2)\big), \qquad \text{for all } x_1, x_2 \in \mathbb{R}. \tag{5}$$

Moreover, if the distribution function F has a bivariate density $f: \mathbb{R}^2 \to [0, \infty)$, the following differential version of Eq. (5) holds:

$$f(x_1, x_2) = c\big(F_1(x_1), F_2(x_2)\big) f_1(x_1) f_2(x_2), \qquad \text{for all } x_1, x_2 \in \mathbb{R}, \tag{6}$$

where $f_i: \mathbb{R} \to [0, \infty)$ is the univariate probability density corresponding to the distribution function F_i, $i = 1, 2$, and the so-called copula density $c: [0,1] \to [0, \infty)$ is the bivariate probability density corresponding to C. In particular, if C is differentiable, then the copula density c is given by

$$c(u_1, u_2) = \frac{\partial^2 C(u_1, u_2)}{\partial u_1 \partial u_2}, \qquad \text{for all } u_1, u_2 \in [0,1]. \tag{7}$$

In the present paper, a modeling approach in the sense of Eq. (6), using parametric fits for the univariate probability densities f_1, f_2 and the copula density c, is employed for the bivariate probability density of sphericity and volume of particles computed from tomographic image data as well as for samples drawn from the models X_Q and X_G introduced in Section 4.1. The resulting parametric fits for the bivariate probability densities are denoted by f_Q^{tomo}, f_G^{tomo}, f_Q^{model} and f_G^{model}, for tomographic image data of quartz and glass particles, and for the realizations of X_Q and X_G, respectively. The specific parametric families used to model the bivariate probability densities for each case are provided in the appendix, see Tables A4 and A5. Furthermore, visualizations of these densities are given in Figure A1.

4.3.2. *Rejection sampling.* In order to reproduce the densities f_Q^{tomo} and f_G^{tomo} computed from from tomographic image data by means of samples drawn from the combined model $X_{Q,G}$, we deploy rejection sampling [31]. In general, rejection sampling is one of the most accessible and widely applicable methods to sample from a desired target distribution, given that it is possible to sample from a so-called source distribution that is reasonably close. The idea is to sample repeatedly from the source distribution, and either reject or accept the obtained realization with a probability that depends on the ratio of two probability densities.

More precisely, let $f, g \colon \mathbb{R}^2 \to \mathbb{R}$ be the probability densities of the source and target distribution, respectively. At first, a normalization constant $M \geq 1$ is chosen such that

$$\frac{g(x)}{Mf(x)} \leq 1, \tag{8}$$

for all $x \in \mathbb{R}^2$. Note that for the existence of such an M, it is necessary that the support of the target probability density g is contained within the support of the source probability density f, *i.e.*, if for some $x \in \mathbb{R}$ it holds that $f(x) = 0$, then also $g(x) = 0$. In our case, this condition is even sufficient, as the densities which appear in our modeling approach are all bounded with bounded support. The procedure of drawing a sample from the probability density g is then given as follows:

(i) Draw a sample $x \in \mathbb{R}^2$ from the source probability density f.

(ii) Draw a uniformly distributed pseudo-random number u between 0 and 1.

(iii) If $u \leq \frac{g(x)}{Mf(x)}$, then accept x. Otherwise, reject x and repeat from (i).

Note that the probability of acceptance is given by $1/M$, see Chapter 3 in [31]. Hence, it is desirable to choose M as small as possible such that Condition (8) is satisfied, *i.e.*, $M = \sup_{x \in \mathbb{R}} g(x)/f(x)$. As already mentioned above, in our case a feasible choice of M is possible, because all parametric univariate densities and bivariate copula densities used for fitting are bounded.
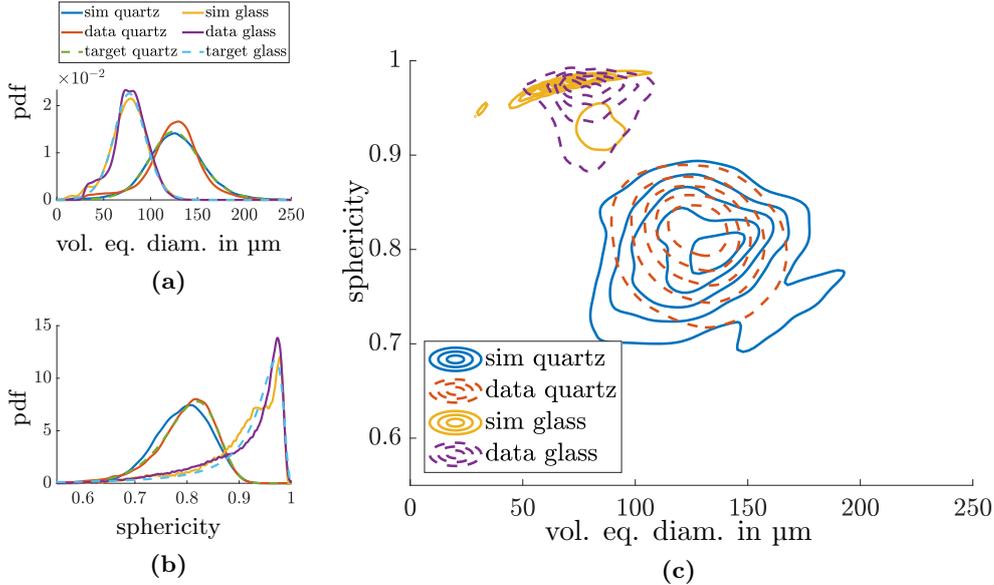
FIGURE 6. Number-weighted distributions of volume-equivalent diameter and sphericity of measured and simulated particles of both glass and quartz particle systems: the univariate probability densities of (a) volume-equivalent diameter and (b) sphericity, and (c) their bivariate joint probability density. Simulated particles are drawn from the combined model $X_{Q,G}$ using copula-based rejection sampling. The parameteric fits f_Q^{tomo} and f_G^{tomo} used as the target distributions of the rejection sampling are indicated by dashed lines. For tomographic image data, 155890 glass and 57977 quartz particles have been considered. For simulated data, 50000 particles were generated in each case.

We apply the rejection sampling approach described above as follows. Since we draw samples from the combined model $X_{Q,G}$, *i.e.*, from both models X_Q and X_G with equal likelihood, the

density f of the source distribution is given by a mixture of the parametric densities fitted in Section 4.3.1, which means that

$$f(x) = \frac{1}{2}\big(f_{\mathrm{Q}}^{\mathrm{model}}(x) + f_{\mathrm{G}}^{\mathrm{model}}(x)\big), \tag{9}$$

for all $x \in \mathbb{R}^2$. After having drawn a virtual particle from $X_{\mathrm{Q,G}}$, we compute its sphericity and volume, which constitutes the sampling of a vector $x \in \mathbb{R}^2$ as in step (i) of the rejection sampling algorithm stated above. The normalization constant M was heuristically set to 20, which did not result in any violations of Condition (8). The target density g is then chosen as the parametric fit of the joint density of volume and sphericity computed from tomographic image data of either the quartz or glass particle system, given by $f_{\mathrm{Q}}^{\mathrm{tomo}}$ and $f_{\mathrm{G}}^{\mathrm{tomo}}$, respectively. Analogously to Figure 5, the univariate and joint probability densities of (number-weighted) particle volume and sphericity obtained for simulated particles via rejection sampling are visualized in Figure 6. While there are still differences between the bivariate densities of glass particles simulated with rejection sampling and glass particles observed in tomographic image data, we can see a significant improvement in comparison to Figure 5, especially for the more aspherical quartz particles.

5. Virtual filter cake model

With the single particle model described in the previous section at hand, we can now simulate an arbitrary number of particles that meet a given joint distribution of volume and sphericity. The next step is to arrange these particles in a cuboidal reference volume, which stands for a representative part of a filter cake. The size of this volume is determined by the desired porosity of the resulting particle system, *i.e.*, the volume fraction of the pore space. At first, a fixed number $n > 1$ of particles is drawn from $X_{\mathrm{Q,G}}$ using rejection sampling as described in Section 4.3.2. The total volume of these particles can be used to determine the necessary size of the cuboidal sampling window, so that the final porosity of the packed particle system matches the one observed in tomographic image data. That is, for a given target porosity $\varepsilon \in [0, 1]$, we obtain a side length of the cuboidal sampling window of

$$a = \bigg(\sum_{i=1}^{n} \frac{V_i}{1 - \varepsilon}\bigg)^{1/3}, \tag{10}$$

where V_i is the volume of the i-th particle drawn from the combined model $X_{\mathrm{Q,G}}$. In order to spatially arrange the n generated particles, placeholder spheres with volume-equivalent radii are initially positioned by sampling their centre coordinates uniformly within the cuboidal volume. Then, a force-biased algorithm [32, 33] is applied to reduce the overlap of the sphere system as far as possible. This is done by iteratively moving each sphere according to a force that is determined by the overlap with neighboring spheres. If a sphere is free of overlaps, it will not be moved any further, unless the movement of other spheres introduces an overlap in a later iteration. As the porosities which have been observed in the samples considered in this paper are above 0.35, we obtain an overlap-free sphere system within reasonable runtime of the algorithm. Finally, each placeholder sphere is replaced by a discretized version of the corresponding originally simulated particle. For this, the particles drawn from the combined model $X_{\mathrm{Q,G}}$ are discretized along the standard Cartesian coordinate system. Note that since the mixed Gaussian random fields modelling the outer particle shapes are isotropic, their distribution is invariant under rotation. Thus, discretizing the particles along a differently chosen orientation would not affect the distribution of the radius functions. However, this replacement of spheres by non-spherical particles generated by $X_{\mathrm{Q,G}}$ can introduce slight overlaps within the particle system. A more detailed and quantitative discussion of this effect is given in Section 6.2 below.

6. Results and discussion

6.1. **Geometric descriptors of pore space.** We validate our model for virtual filter cake structures with respect to various geometric descriptors of the pore space. For this, we simulate 10 virtual filter cakes for both systems of quartz and glass particles, consisting of 5000 particles each as described in Sections 4 and 5, see Figure 7. From the binary 3D images resulting from this simulation, a layer of 20 voxels thickness (60 µm) is removed from each side for subsequent computation of geometric descriptors. The average size of the sampling windows obtained in this way is 511^3 and 312^3 voxels for quartz and glass particles, respectively.
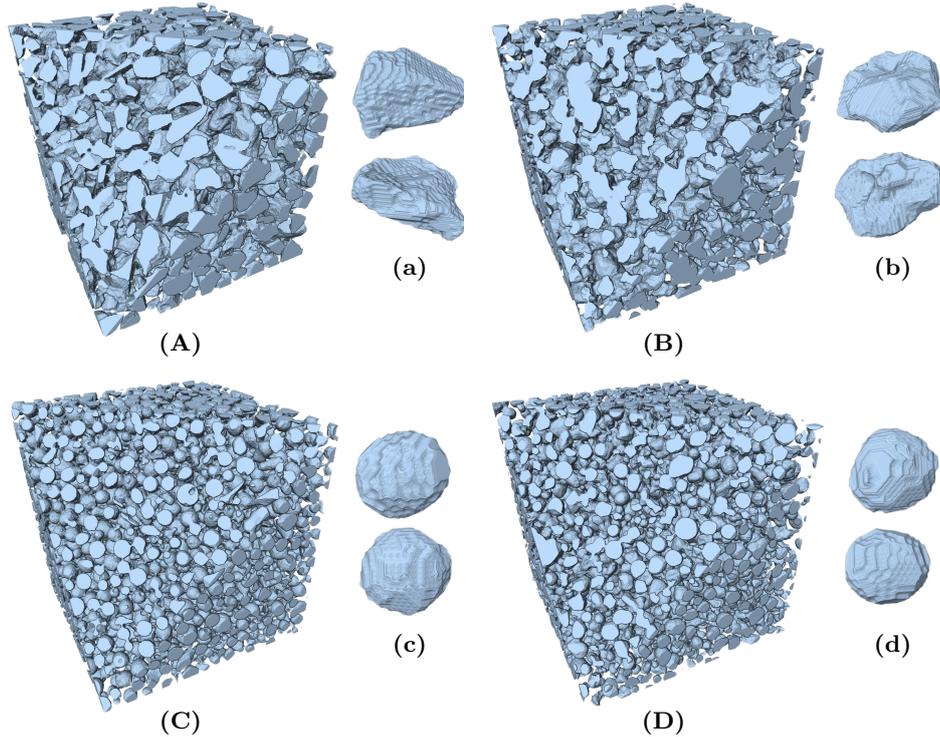
FIGURE 7. 3D renderings of measured and simulated filter cake structures. Top row: filter cakes based on measured (A) and simulated (B) quartz particles. Bottom row: filter cakes based on measured (C) and simulated (D) glass particles. For visualization purposes, the renderings in subfigures (A) and (C) only show small cutouts of the available tomographic image data. Furthermore, two different perspectives of single particles are shown for measured (a) and simulated (b) quartz particles, as well as measured (c) and simulated (d) glass particles.

The simulated filter cakes are then compared to tomographic image data of the corresponding particle system. The geometric descriptors which we consider for this comparison have been chosen as they showed to significantly influence effective transport-related properties in functional materials [45].

First, we consider the porosity ε, being the volume fraction of the pore space. This geometric descriptor is computed by simply counting the number of voxels that belong to the pore space and dividing it by the total number of voxels in the sampling window. We also consider the specific surface area S_V, *i.e.*, the surface area per unit volume, which is estimated based on voxelized image data using the algorithm proposed in [46], which considers a weighted $2 \times 2 \times 2$ neighborhood of each voxel. This quantity is indirectly used for model calibration, as our rejection sampling approach is designed to reproduce the desired sphericity distribution among individual particles, and the sphericity of particles is closely related with the surface area of

pore space. These two descriptors, ε and S_V, of the filter cake structure are primarily influenced by shape factors of individual particles, namely volume and surface area, which are fitted by design of the single particle model. However, the remaining geometric descriptors that we use for model validation are not related to the fitting procedure of the single particle model, which makes them more suitable for validation purposes.

One of these descriptors is the mean geodesic tortuosity τ, which quantifies the length of shortest paths through the material within the pore space in comparison to the thickness of the material. We consider paths along the z-direction, as this aligns with the natural flow occurring in the pore space during filter cake formation. For each pore space voxel of the first slice of the tomographic image data in z-direction, we compute the length of the shortest path through the pore space and divide it by the height of the tomographic stack in z-direction. The mean geodesic tortuosity τ is then estimated by the sample mean of these ratios. We stress that there are many different notions of tortuosity which differ in the types of paths that they consider. For an overview of various notions of tortuosity, we refer the reader to [47].

Next we consider the so-called continuous pore size distribution defined as in [48], *i.e.,* the function CPSD: $[0, \infty) \to [0, 1]$, which is based on morphological opening [39, 49]. For each $r \geq 0$, the value CPSD(r) is estimated by the fraction of voxels of the pore space which can be covered by spheres that are fully contained within the pore space. Formally, let $P \subset W$ denote the subset of voxels belonging to the pore space within our sampling window $W \subset \mathbb{Z}^3$. Furthermore, for any subset $A \subset W$, by $|A|$ we denote the cardinality of A. The continuous pore size distribution is then given by CPSD$(0) = 1$ and

$$\text{CPSD}(r) = \frac{\left| \left(P \ominus B(o, r) \right) \oplus \left(B(o, r) \cap \mathbb{Z}^3 \right) \right|}{|P|}, \tag{11}$$

for any $r > 0$, where $\ominus$ and $\oplus$ denote morphological erosion and dilation [39, 49], respectively, and $B(o, r) \subset \mathbb{R}^3$ is the open ball in the three-dimensional Euclidean space $\mathbb{R}^3$ with radius r centered at the origin $o \in \mathbb{R}^3$. An efficient implementation for the computation of CPSD(r) for $r > 0$ is given by means of the Euclidean distance transform [39, 50].

Another geometric descriptor of the pore space is the so-called simulated mercury intrusion porosimetry. Here, similar to CPSD, we consider a function MIP: $[0, \infty) \to [0, 1]$, which is related with covering the pore space by spheres of varying radii that are completely contained within the pore space. However, we now consider spheres that intrude the pore space in W from a predefined direction. Thus, in contrast to CPSD(r), a position within the pore space is considered to be coverable in the context of MIP(r) if it is reachable by a sphere of radius $r \geq 0$ from the predefined direction. In this way, we can account for and quantify bottleneck effects within the pore space. Formally, we can write

$$\text{MIP}(r) = \frac{\left| \mathcal{C}\left(P \ominus B(o, r), H_0 \right) \oplus \left(B(o, r) \cap \mathbb{Z}^3 \right) \right|}{|P|} \tag{12}$$

for any $r \geq 0$, where for any set of voxels $A \subset W$, the set $\mathcal{C}(A, H_0) \subset A$ denotes the subset of voxels within A that are reachable by a path that is contained in A and starts in a predefined planar set of voxels $H_0 \subset A$. In our case, the planar voxel set H_0 is chosen to be aligned with the (x, y)-plane, so that the direction of simulated mercury intrusion porosimetry is aligned with the natural flow occurring in the pore space during filter cake formation. However, for the tomographic image data considered in this paper, different choices of H_0 lead to nearly identical resulting MIP-curves which is a strong indication of isotropy. The computation of the set $\mathcal{C}(A, H_0)$ is implemented by means of the Hoshen-Kopelman algorithm [51].

Finally, we consider the constrictivity β, which is a geometric descriptor of pore space that quantifies the strength of bottleneck effects, using information of the functions CPSD and MIP explained above. In particular, we consider two characteristic radii, denoted by $r_{\max}$ and $r_{\min}$, which are defined as solutions of the equations CPSD$(r_{\max}) = 1/2$ and MIP$(r_{\min}) = 1/2$,

respectively. The choice of the threshold $1/2$ in these equations allows us to interpret $r_{\min}$ and $r_{\max}$ as median pore radii. The constrictivity β of the pore space is then given by $\beta = \left(\frac{r_{\min}}{r_{\max}}\right)^2$ [48]. Note that, due to the additional reachability condition considered in Eq. (12), compared to Eq. (11), it holds that $\mathrm{MIP}(r) \leq \mathrm{CPSD}(r)$ for each $r \geq 0$. Hence, $r_{\min} \leq r_{\max}$ and, consequently, $\beta \in [0, 1]$. A constrictivity β close to 1 indicates that the additional reachability requirement in Eq. (12) has little influence on the resulting MIP-curve in comparison to the CPSD-curve, which implies that there are only few or no constrictions within the pore space. On the other hand, a constrictivity β near 0 indicates significant bottlenecks along the transport paths through the pore space. For more information on the definition and estimation of mean geodesic tortuosity and constrictivity within the formal framework of stationary random closed sets, see [52].

6.2. **Model validation.** We now validate both the single particle model introduced in Section 4.1 as well as the filter cake model stated in Section 5.

6.2.1. *Single particle model.* Gerometric descriptors of virtual particles drawn from the single particle models X_Q and X_G introduced in Section 4.1 (*i.e.*, prior to the introduction of the rejection sampling scheme considered in Section 4.3) are compared with those of particles observed in tomographic image data, see Figure 5. While the nearly spherical glass particles are well presented by the model both in size and shape, the more aspherical quartz particles are not accurately reproduced, as the sphericity distribution of simulated quartz particles does not align well with the sphericity distribution of measured quartz particles.

Therefore, we introduced a method for combining the single particle model described in Section 4.1 with a copula-based rejection sampling scheme, see Section 4.3. This allows us to reject virtual particles with a probability depending on their size and shape, so that the resulting joint distribution of volume and sphericity matches the corresponding distribution obtained for particles observed in tomographic image data. In view of the results presented in Figure 6, we can claim that this approach is successful, since the joint distribution of volume and sphericity of measured particles is now accurately reproduced by virtual particles generated by means of rejection sampling. Note that these fits have been achieved with a single combined model that uses information of both data sets for quarz and glass particles. However, the plots shown in Figures 5c and 6c provide just a qualitative comparison of the involved densities. For more detailed information, we refer to the heat maps provided in Figure A2 of the appendix.

6.2.2. *Filter cake model.* To validate the structure of simulated filter cakes, we determined the distribution of various geometric descriptors stated in Section 6.1, both for filter cakes observed in tomographic image data as well as for simulated filter cakes, see Figure 8. For most of these descriptors, there is no direct connection to the shape and size of individual particles, so that it is a priori unclear whether they would show good agreement for virtually generated filter cake structures and those observed in tomographic image data. Nevertheless, Figure 8 shows that all geometric descriptors of pore space considered in this paper are accurately reproduced by the model. Moreover, the morphologies of the pore space in filter cakes consisting of quartz and glass particles, respectively, differ significantly from each other, both in the case of measured and simulated filter cakes. Since the packing has been performed in the same way for simulated quartz and glass particles, we can conclude that there is a clear influence of the size and shape distribution of individual particles on the 3D morphology of the pore space of corresponding filter cakes, as was also found in other experimental and numerical studies [53–55].

As stated in Section 5, the virtual quartz and glass particles generated by the rejection sampling scheme are used to simulate virtual filter cake structures by means of the force-bias algorithm, where only the particle size distribution and the porosity are taken into account in this packing algorithm. Thus, the influence of particle shape on the packed particle system is neglected. More precisely, the forced-bias algorithm uses volume-equivalent spheres for

collision/overlap detection. As a result, a system of spherical particles is initially generated. Then, in a next step, the placeholder spheres are substituted by individually shaped particles, which does not introduce preferential particle orientations, due to the isotropy of the underlying Gaussian random field $X_{Q,G}$, see Section 4.1. However, according to [24], the shapes of individual particles and their distribution have a decisive influence on the 3D morphology of filter cakes, which increases strongly with increasing deviation from the spherical shape, *i.e.*, with decreasing sphericity. With the measured sphericities between 0.7 and 1.0 for the glass particles and between 0.6 and 0.9 for quartz, see Figure 6b, the deviations from the spherical shape are relatively small, and hence the usage of spherical particles in the packing process seems to be reasonable. Moreover, the good fit of the geometric pore space descriptors shown in Figure 8 justifies this simplifying assumption.
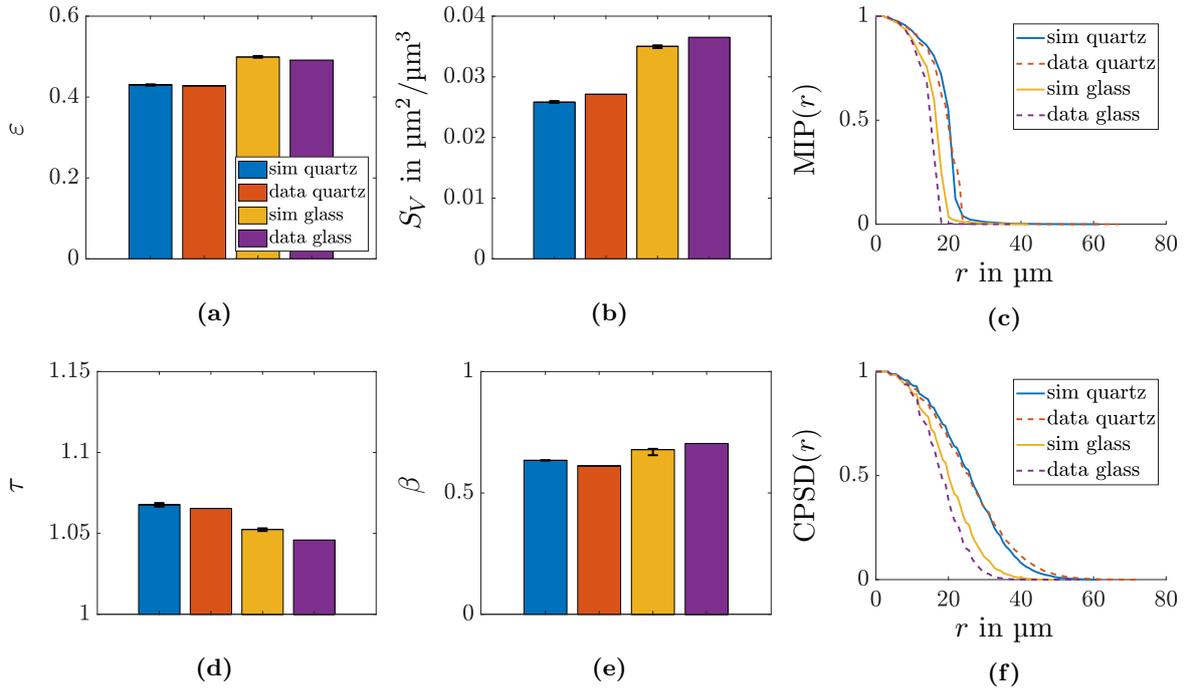
FIGURE 8. Geometric pore space descriptors of measured and virtual filter cakes for both glass and quartz particles, where 10 virtual filter cake structures consisting of 5000 simulated particles each were generated for both particle systems: (a) porosity ε, (b) surface area per unit volume S_V, (c) simulated mercury intrusion porosimetry, (d) mean geodesic tortuosity τ, (e) constrictivity β, and (f) continuous pore size distribution.

Interestingly, the deviation between the CPSD curves of simulated and measured filter cakes, as seen in Figure 8f, is larger for glass than for quartz particles, which is against the intuition that the clearly non-spherical shape of quartz particles neglected in the packing process could introduce errors that are not seen for the spherical glass particles. We believe that this effect could be due to a slight clustering in the artificial filter cake structures consisting of glass particles. As the porosity of the glass filter cake is higher, the packing is more loose and allows for some maneuvering room amongst the particles. However, the placeholder spheres are only moved as far as necessary to receive on overlap-free system, see Section 5, which can lead to clustered particles on the one hand, and some larger pores on the other hand. This effect can be seen in Figure 8f, where the simulated glass structure shows a slightly increased fraction of large pores ($\geq 35\,\mu\mathrm{m}$), and a slightly decreased amount of smaller pores ($\approx 20\,\mu\mathrm{m}$) in comparison to

the tomographic image data of a measured glass filter cake. Regarding the clustering of glass particles mentioned above we refer to the 2D slices of both simulated and measured tomographic image data provided in Figures A3a–A3c of the appendix.

6.2.3. *Analysis of particle overlap.* As mentioned above, the deployed sphere-packing algorithm solely minimizes the overlap of the placeholder spheres, whereas the replacement of these spheres with individually shaped particles may introduce slight additional overlaps, which we now aim to discuss in more detail. We can already see this effect in Figures 8a and 8b. The porosities of simulated filter cakes are slightly higher than those of measured ones, which indicates a small amount of lost particle volume. The same is true for the surface area per unit volume, where some surface area is lost whenever two particles overlap. In order to obtain a more quantitative result on this effect, we perform the same particle-discrete segmentation procedure that was used for the segmentation of the tomographic image data in Section 3 on the binary image data of virtually generated filter cake structures. The resulting distributions of (number-weighted) volume and sphericity of the segmented particles are visualized in Figure 9.
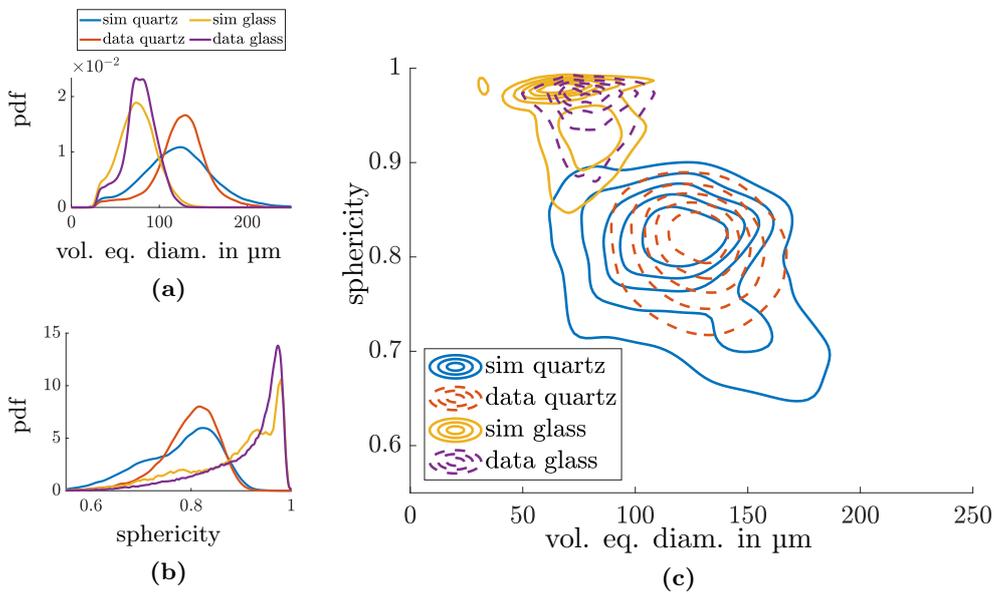
FIGURE 9. Number-weighted distributions of volume-equivalent diameter and sphericity of particles in measured and virtual filter cakes of both glass and quartz particle systems: the univariate probability densities of (a) volume equivalent diameter and (b) sphericity, and (c) their bivariate joint probability density. For simulated data, 10 virtual filter cake structures consisting of 5000 particles each were generated for both particle systems. Based on the resulting binary image data, a particle-discrete segmentation was performed in the same manner as for tomographic image data.

In Figure 9a, we observe that the number-weighted volume distributions of both quartz and glass particles within simulated filter cakes are shifted towards lower values in comparison to particles observed in tomographic image data. This shift is likely contributed to lost volume due to overlaps between particles after the force-bias algorithm. Since the probability densities in Figure 9a are number-weighted and not volume-weighted, the large deviation of the densities for volumes close to 0 contributes only to a relatively small volume of particles within the considered filter cakes. In addition, there is a slight shift towards lower values of sphericity for the quartz particle system. This is likely caused by the segmentation procedure, as the watershed algorithm

detects boundaries between touching particles as flat faces, which can reduce the sphericity of particles if even a slight overlap is present, while it would only rarely increase the sphericity of involved particles. Although the introduced overlaps have visible effect on the distribution of volume and sphericity of particles after the deployment of the watershed algorithm, we observe that the impact is only minor and we could not establish that the pore morphology of the resulting filter cake structure is affected in a significant way.

7. Conclusion

The model presented in this paper for individual particle sizes and shapes is based on a spherical harmonics approach, which is combined with a copula-based rejection sampling scheme. In this way, typical limitations in the shape of particles generated by means of spherical harmonics functions can be overcome. Moreover, by modifying the desired target distribution within the rejection sampling scheme, particles with various sizes and shapes can be generated by simply modifying the parameters of their copula-based bivariate probability density, without the need to re-calibrate the model to tomographic image data. Virtual particles generated in this way are packed into filter cake structures using a force-biased algorithm. The 3D morphology of virtual filter cakes is validated by comparing their geometric pore space descriptors with corresponding descriptors determined for filter cakes observed in tomographic image data. The presented approach sets up a framework for a large scale simulation study in order to quantify the influence of individual particle shapes on the pore space morphology of the resulting filter cakes. This provides the necessary understanding to further investigate the influence of individual particle shapes on effective properties of filter cakes, which ultimately influence the process properties of cake filtration and downstream processes like de-watering and washing of the porous structure. Thus, in future research, the copula-based rejection sampling scheme developed in the present paper will be used to generate particles that follow a wide range of joint distributions of volume and sphericity.

Data availability

Complete raw data (reconstructed TIFF stacks) of all filter cakes considered in this study is available by open access via the OpARA online repository of TU Dresden and TU Bergakademie Freiberg (http://dx.doi.org/10.25532/OPARA-292).

Acknowledgements

References

[1] S. E. Woodard and R. F. Wukasch. A hydrolysis/thickening/filtration process for the treatment of waste activated sludge. *Water Science and Technology*, 30:29–38, 1994.

[2] A. Sonune and R. Ghate. Developments in wastewater treatment methods. *Desalination*, 167:55–63, 2004.

[3] C. Wang, D. Harbottle, Q. Liu, and Z. Xu. Current state of fine mineral tailings treatment: A critical review on theory and practice. *Minerals Engineering*, 58:113–131, 2014.

[4] B. F. Ruth, G. H. Montillon, and R. E. Montonna. Studies in filtration - I. Critical analysis of filtration theory. *Industrial & Engineering Chemistry*, 25:76–82, 1933.

[5] P. C. Carman. Fluid flow through granular beds. *Chemical Engineering Research and Design*, 75:S32–S48, 1997.

[6] J. Sauter. *Die Grössenbestimmung der im Gemischnebel von Verbrennungskraftmaschinen vorhandenen Brennstoffteilchen: Mitteilung aus d. Laboratorium f. techn. Physik d. Techn. Hochschule München.* Forschungsarbeiten auf dem Gebiete des Ingenieurwesens. VDI-Verlag, 1926.

[7] P. C. Carman. Some physical aspects of water flow in porous media. *Discussions of the Faraday Society*, 3:72–77, 1948.

[8] C. Tien and B. V. Ramarao. Can filter cake porosity be estimated based on the Kozeny–Carman equation? *Powder Technology*, 237:233–240, 2013.

[9] D. J. Condie, M. Hinkel, and C. J. Veal. Modelling the vacuum filtration of fine coal. *Filtration & Separation*, 33:825–834, 1996.

[10] J. P. Du Plessis and J. H. Masliyah. Mathematical modelling of flow through consolidated isotropic porous media. *Transport in Porous Media*, 3:145–161, 1988.

[11] S. Suo, M. Liu, and Y. Gan. Modelling imbibition processes in heterogeneous porous media. *Transport in Porous Media*, 126:615–631, 2019.

[12] P. Iassonov, T. Gebrenegus, and M. Tuller. Segmentation of X-ray computed tomography images of porous materials: A crucial step for characterization and quantitative analysis of pore structures. *Water Resources Research*, 45:W09415, 2009.

[13] Ç. Hoşten and K. V. S. Sastry. Empirical correlations for the prediction of cake dewatering characteristics. *Minerals Engineering*, 2:111–119, 1989.

[14] E. Löwer, F. Pfaff, T. Leißner, and U.A. Peuker. Neighborhood relationships of widely distributed and irregularly shaped particles in partially dewatered filter cakes. *Transport in Porous Media*, 138:201–224, 2021.

[15] K. Hutton, N. Mitchell, and P. J. Frawley. Particle size distribution reconstruction: The moment surface method. *Powder Technology*, 222:8–14, 2012.

[16] E. Löwer, T. H. Pham, T. Leißner, and U. A. Peuker. Study on the influence of solids volume fraction on filter cake structures using micro tomography. *Powder Technology*, 363:286–299, 2020.

[17] R. G. Shepherd. Correlations of permeability and grain size. *Groundwater*, 27:633–638, 1989.

[18] M. N. Panda and L. W. Lake. Estimation of single-phase permeability from parameters of particle-size distribution. *AAPG Bulletin*, 78:1028–1039, 1994.

[19] G. Mavko and A. Nur. The effect of a percolation threshold in the Kozeny-Carman relation. *Geophysics*, 62:1480–1482, 1997.

[20] H. Pape, C. Clauser, and J. Iffland. Permeability prediction for reservoir sandstones and basement rocks based on fractal pore space geometry. *SEG Technical Program Expanded Abstracts 1998*, pages 1032–1035, 1998.

[21] F. Civan. Scale effect on porosity and permeability: Kinetics, model, and correlation. *AIChE Journal*, 47:271–287, 2001.

[22] A. Costa. Permeability-porosity relationship: A reexamination of the Kozeny-Carman equation based on a fractal pore-space geometry assumption. *Geophysical Research Letters*, 33, 2006.

[23] P. Xu and B. Yu. Developing a new form of permeability and Kozeny-Carman constant for homogeneous porous media by means of fractal geometry. *Advances in Water Resources*, 31:74–81, 2008.

[24] H. Rumpf and A. R. Gupte. Einflüsse der Porosität und Korngrößenverteilung im Widerstandsgesetz der Porenströmung. *Chemie Ingenieur Technik*, 43:367–375, 1971.

[25] B. Prifling, D. Westhoff, D. Schmidt, H. Markötter, I. Manke, V. Knoblauch, and V. Schmidt. Parametric microstructure modeling of compressed cathode materials for Li-ion batteries. *Computational Materials Science*, 169:109083, 2019.

[26] S. N. Chiu, D. Stoyan, W. S. Kendall, and J. Mecke. *Stochastic Geometry and its Applications.* J. Wiley & Sons, 3rd edition, 2013.

[27] V. Schmidt, editor. *Stochastic Geometry, Spatial Statistics and Random Fields: Models and Algorithms*. Springer, 2015.

[28] A. Lang and C. Schwab. Isotropic Gaussian random fields on the sphere: Regularity, fast simulation and stochastic partial differential equations. *The Annals of Applied Probability*, 25:3047–3094, 2015.

[29] O. Furat, L. Petrich, D. P. Finegan, D. Diercks, F. Usseglio-Viretta, K. Smith, and V. Schmidt. Artificial generation of representative single Li-ion electrode particle architectures from microscopy data. *npj Computational Materials*, 7:105, 2021.

[30] H. Joe. *Dependence Modeling with Copulas*. CRC Press, 2014.

[31] L. Martino, D. Luengo, and J. Míguez. *Independent Random Sampling Methods*. Springer, 2018.

[32] J. Mościński, M. Bargieł, Z. A. Rycerz, and P. W. M. Jakobs. The force-biased algorithm for the irregular close packing of equal hard spheres. *Molecular Simulation*, 3:201–212, 1989.

[33] A. Bezrukov, M. Bargieł, and D. Stoyan. Statistical analysis of simulated random packings of spheres. *Particle & Particle Systems Characterization*, 19:111–118, 2002.

[34] The Association of German Engineers. Mechanical solid-liquid separation by cake filtration determination of filter cake resistance. Beuth Verlag GmbH, 2010. ICS 71.040.40.

[35] L. Gerward, N. Guilbert, K. B. Jensen, and H. Levring. WinXCom—a program for calculating X-ray attenuation coefficients. *Radiation Physics and Chemistry*, 71:653–654, 2004.

[36] G. T. Herman. *Fundamentals of Computerized Tomography: Image Reconstruction from Projections*. Springer, 2009.

[37] C. A. Glasbey. An analysis of histogram-based thresholding algorithms. *CVGIP: Graphical Models and Image Processing*, 55:532–537, 1993.

[38] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. White, V. Hartenstein, K. Eliceiri, P. Tomancak, and A. Cardona. Fiji: An open-source platform for biological-image analysis. *Nature Methods*, 9:676–82, 2012.

[39] P. Soille. *Morphological Image Analysis: Principles and Applications*. Springer, 2003.

[40] L. Vincent. Morphological grayscale reconstruction: Definition, efficient algorithm and applications in image analysis. In *Proceedings of 1992 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 633–635. IEEE Computer Society, 1992.

[41] J. Zheng and R. D. Hryciw. Segmentation of contacting soil particles in images by modified watershed analysis. *Computers and Geotechnics*, 73:142–152, 2016.

[42] P. Gräfensteiner, J. Friebel, L. Ditscherlein, O. Furat, U. A. Peuker, and V. Schmidt. An AFM-based approach for quantification of guest particle deformation during mechano-fusion. *Powder Technology*, 434:119293, 2024.

[43] D. Marinucci and G. Peccati. *Random Fields on the Sphere: Representation, Limit Theorems and Cosmological Applications*. Cambridge University Press, 2011.

[44] H. Joe. *Multivariate Models and Dependence Concepts*. Chapman & Hall, 1997.

[45] L. Holzer, D. Wiedenmann, B. Münch, L. Keller, M. Prestat, P. Gasser, I. Robertson, and B. Grobéty. The influence of constrictivity on the effective transport properties of porous layers in electrolysis and fuel cells. *Journal of Materials Science*, 48:2934–2952, 2013.

[46] J. Ohser and K. Schladitz. *3D Images of Materials Structures: Processing and Analysis*. Wiley-VCH, 2009.

[47] L. Holzer, P. Marmet, M. Fingerle-Straß, A. Wiegmann, M. Neumann, and V. Schmidt. *Tortuosity and Microstructure Effects in Porous Media: Classical Theories, Empirical Data and Modern Methods*. Springer, 2023.

[48] B. Münch and L. Holzer. Contradicting geometrical concepts in pore size analysis attained with electron microscopy and mercury intrusion. *Journal of the American Ceramic Society*,

91:4059–4067, 2008.

[49] J. Serra. *Image Analysis and Mathematical Morphology.* Academic Press, 1982.

[50] C. R. Maurer, R. Qi, and V. Raghavan. A linear time algorithm for computing exact Euclidean distance transforms of binary images in arbitrary dimensions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25:265–270, 2003.

[51] J. Hoshen and R. Kopelman. Percolation and cluster distribution. I. Cluster multiple labeling technique and critical concentration algorithm. *Physical Review B*, 14:3438–3445, 1976.

[52] M. Neumann, C. Hirsch, J. Staněk, V. Beneš, and V. Schmidt. Estimation of geodesic tortuosity and constrictivity in stationary random closed sets. *Scandinavian Journal of Statistics*, 46:848–884, 2019.

[53] A. Kerimov, G. Mavko, T. Mukerji, J. Dvorkin, and M. A. Al Ibrahim. The influence of convex particles' irregular shape and varying size on porosity, permeability, and elastic bulk modulus of granular porous media: Insights from numerical simulations. *Journal of Geophysical Research: Solid Earth*, 123:10–563, 2018.

[54] A. El-Husseiny, T. Vanorio, and G. Mavko. Predicting porosity of binary mixtures made out of irregular nonspherical particles: Application to natural sediments. *Advanced Powder Technology*, 30:1558–1566, 2019.

[55] E. Rezaei, K. Zeinalzadeh, and B. Ghanbarian. Experimental study of hydraulic properties in grain packs: Effects of particle shape and size distribution. *Contaminant Hydrology*, 243:103918, 2021.

<center>Appendix</center>

In this appendix, we provide additional information on different aspects of the manuscript. In Table A1, filtration parameters for the experimental preparation of the filter cakes are presented. Settings of the XRM measurements used to obtain tomographic image data are shown in Table A2. Parametric families of univariate probability distributions and bi- variate copulas used to model the joint distribution of particle volume and sphericity are given in Table A3. Tables A4 and A5 show the estimated parameters of the model fits obtained for the univariate (marginal) distributions of volume and sphericity, as well as the copula, modeling their dependency structure for quartz and glass particles, respectively. Figure A1 shows a visualization of these parametric fits for the bivariate (joint) probability densities of volume and sphericity. Figure A2 is a more detailed visualization of Figures 5c and 6c of the main manuscript, where the estimated bivariate probability densities of volume equivalent diameter and sphericity are shown for particles observed in tomographic image data (A2a, A2b) and particles simulated separately through the models X_Q and X_G (A2c, A2d), as well as particles drawn from the joint model $X_{Q,G}$ (A2e, A2f). Figure A3 shows examples of 2D slices of simulated and measured 3D image data for both particle systems.

TABLE A1. Overview of filtration parameters for the experimental preparation of filter cakes with mass of solid in suspension m_S, suspension volume fraction φ, filtration pressure Δp and filtration results of the analysis according to VDI 2762-2 [34] for quartz and glass particles. The height specific filter cake resistance is denoted by r_C, while R_M denotes the filter media resistance and RM is the residual moisture after cake filtration.

	quartz (BCS201)	glass (B70)
m_S in g	70	70
φ in m³/m³	0.40	0.45
Δp in kPa	30	30
r_C in 1/m²	$5.7 * 10^{11}$	$1.2 * 10^{12}$
R_M in 1/m	$6.5 * 10^9$	$2.6 * 10^{10}$
RM in kg/kg	0.14	0.11

TABLE A2. Parameter settings for the XRM measurements.

	quartz (BCS201)	glass (B70)
field of view	4 x 4 cm²	4 x 4 cm²
sample size	5 x 5 x 12.4 mm³	5 x 5 x 12.9 mm³
source filter (ZEISS standard)	LE4	LE5
acceleration voltage / power	50 kV/4 W	50 kV/4 W
optical magnification	4x	4x
pixel size	3.98 μm	4.00 μm
exposure time	4 s	8 s
projections	2001(360°)	2001(360°)
camera binning	2	2

TABLE A3. Parametric families of univariate probability distributions and bivariate copulas used to model the joint distribution of particle volume and sphericity, where $\mathbb{1}_C$ denotes the indicator of a set $C \subset \mathbb{R}$, i.e., $\mathbb{1}_C(x) = 1$ if $x \in C$ and $\mathbb{1}_C(x) = 0$ if $x \notin C$.

parametric familiy	density	parameters
GEV	$f(x) = (1 + k\frac{x-\mu}{\sigma})^{-\left(1+\frac{1}{k}\right)}$ $\times \exp\left(-(1+k\frac{x-\mu}{\sigma})^{-\frac{1}{k}}\right)\mathbb{1}_{[-1/k,\infty)}(\frac{x-\mu}{\sigma}), \quad x \in \mathbb{R}$	$k \in \mathbb{R}$, $\sigma > 0$, $\mu \in \mathbb{R}$
Weibull	$f(x) = \frac{B}{A}\left(\frac{x}{A}\right)^{B-1}\exp\left(-\left(\frac{x}{A}\right)^B\right)\mathbb{1}_{[0,\infty)}(x), \quad x \in \mathbb{R}$	$A, B > 0$
Clayton	$c(u_1, u_2) = (1+\rho)(u_1 u_2)^{(-1-\rho)}$ $\times(-1 + u_1^{-\rho} + u_2^{-\rho})^{(-2-1/\rho)}, \quad u_1, u_2 \in [0,1]$	$\rho > -1$

TABLE A4. Parametric families of univariate probability distributions and bivariate copulas, together with fitted parameter values, used to model the joint distribution of volume and sphericity of quartz particles, based on tomographic image data and on model realizations, denoted by f_Q^{tomo} and f_Q^{model} in the main text, respectively.

descriptor	parametric family	fitted parameter values
volume (data)	GEV	$k = 0.0438$, $\sigma = 8780.28$, $\mu = 13131.7$
sphericity (data)	Weibull	$A = 0.822$, $B = 17.353$
copula (data)	Clayton	$\rho = 0.1911$
volume (sim)	GEV	$k = 0.0726$, $\sigma = 7745.25$, $\mu = 11078.5$
sphericity (sim)	Weibull	$A = 0.885$, $B = 18.792$
copula (sim)	Clayton	$\rho = 0.0025$

TABLE A5. Parametric families of univariate probability distributions and bivariate copulas, together with fitted parameter values, used to model the joint distribution of volume and sphericity of glass particles, based on tomographic image data and on model realizations, denoted by f_G^{tomo} and f_G^{model} in the main text, respectively.

descriptor	parametric family	fitted parameter values
volume (data)	GEV	$k = 0.0376535$, $\sigma = 138657$, $\mu = 205103$
1 - sphericity (data)	GEV	$k = 0.598497$, $\sigma = 0.0366298$, $\mu = 0.0455061$
copula (data)	Clayton	$\rho = 8.4972 \cdot 10^{-4}$
volume (sim)	GEV	$k = 0.0577825$, $\sigma = 123180$, $\mu = 166568$
1 - sphericity (sim)	GEV	$k = 0.871166$, $\sigma = 0.0167278$, $\mu = 0.0256364$
copula (sim)	Clayton	$\rho = 3.5464 \cdot 10^{-5}$
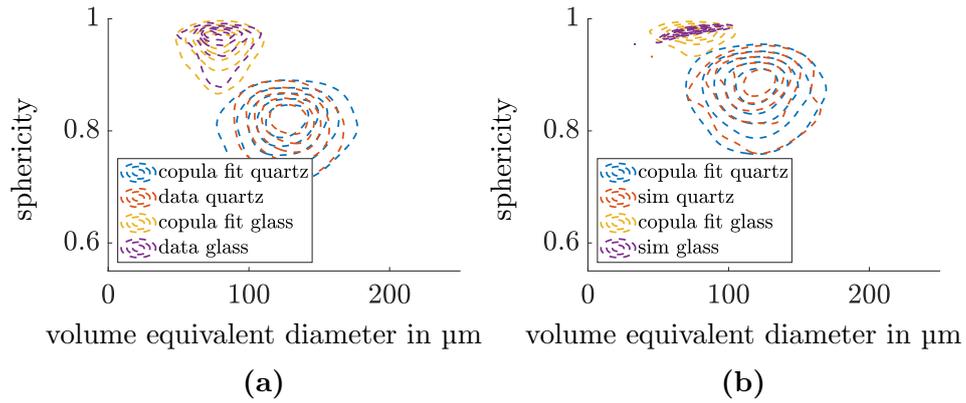
FIGURE A1. Parametric fits of the bivariate joint probability density of volume and sphericity, for quartz and glass particles, observed in (a) tomographic image data and (b) simulated image data.
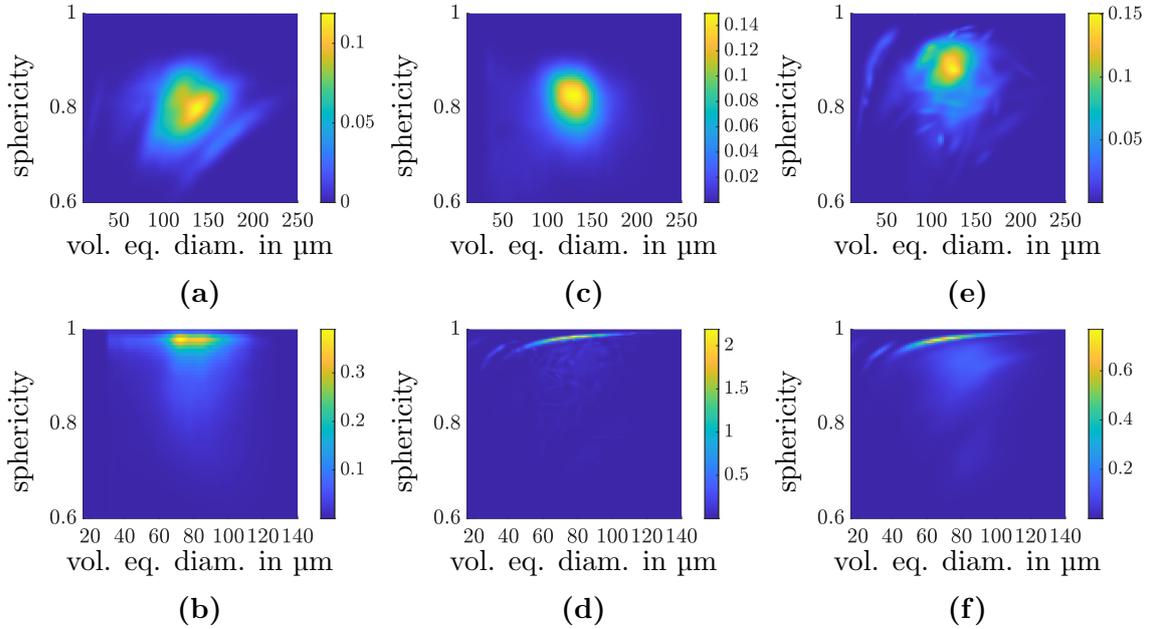
FIGURE A2. Bivariate probability density of volume and sphericity for simulated quartz (top) and glass (bottom) particles drawn from the separately calibrated models X_Q and X_G (a,b), for simulated quartz (top) and glass (bottom) particles drawn from the joint model $X_{Q,G}$, and for quartz (top) and glass (bottom) particles observed in tomographic image data (e,f). For tomographic image data, 57977 quartz and 155890 glass particles were considered. For simulated data, 50000 particles were generated in each case.
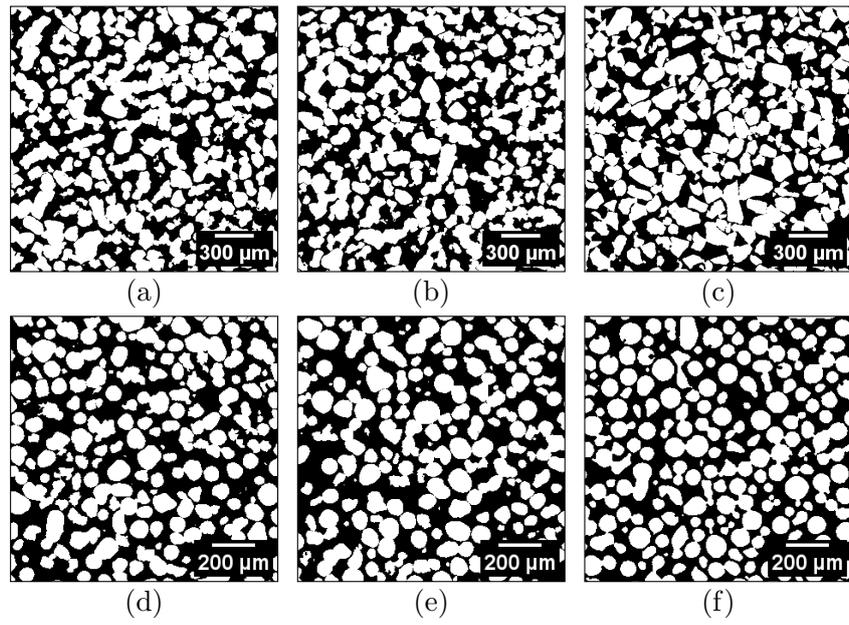
FIGURE A3. Examples of 2D slices of simulated and measured 3D image data for simulated quartz particle filter cakes based on model $X_{Q,G}$ (a,b), measured quartz particle filter cake (c), simulated glass particle filter cakes based on model $X_{Q,G}$ (d,e), and measured glass particle filter cake (f).