Morphology evolution of crystal populations: Modeling and observation analysis

Christian Borchert[a], Kai Sundmacher[a,b]*

[a] Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstrasse 1, D-39106 Magdeburg, Germany
[b] Otto von Guericke University, Process Systems Engineering, PSF410, D-39106 Magdeburg, Germany

ABSTRACT

The joint analysis of modeling and observation of crystallization systems with regard to shape is taken out. We start with the modeling and numerical solution of the balance equations for a seeded batch crystallizer accounting for crystal shape distributions. For this, a meshing algorithm is used to establish appropriate origins for characteristic curves of the population balance. The calculated number density evolution is sampled and the resulting representative crystal shapes are photographed. The orientation of the crystal relative to the photoplane is modeled as a stochastic process. Therefore, the projection cannot be related to the actual 3D shape in a trivial way. We design an estimation scheme and evaluate its reliability for some test populations. Finally, the number density evolution of a whole simulated experiment is observed at different sampling rates. We show that the underlying growth parameters can be reestimated from the observed data in a reliable way for which a relatively small number of samples is required.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

About 60% of all products produced by major chemical companies are delivered as solids, among them many crystalline materials (Wintermantel, 1999). Virtually all pharmaceutical production processes involve a crystallization step and most active pharmaceutical ingredients are administered in a crystal-line form (Variankaval et al., 2008). Crystalline pharmaceuticals, agrochemicals, cosmetics and fine chemicals are high value-added products for which crystal shape is an important quality factor. Though it is well known that properties of dispersed phase products are strongly linked to their shape, process systems engineering research was so far focused on particle size and size distributions and only during the last years efforts have been started to include quantitative measures for shape and shape distributions, e.g. Bajinca et al. (2010), Borchert et al. (2009b), Briesen (2006), Chakraborty et al. (2010), Kempkes et al. (2010b) and Ma et al. (2002). Examples where shape has been of interest - not necessarily in quantitative terms or along with model devel-opment - range over the whole palette from bulk chemical products (Aquilano et al., 2009; Liu et al., 1995) over nanoparticle applications (Barnard, 2009; Chemseddine and Moritz, 1999) to catalysis (Selloni, 2008; Yang et al., 2008; Yang and Liu, 2009; Yi

et al., 2009). An overview on crystal shape engineering and recent advances with a special focus on solution crystallization has been published by Doherty and coworkers (Lovette et al., 2008).

The understanding of a complex process requires a realistic and thus physically interpretable model. The dynamics of dis-persed phase processes are at best captured with a population balance equation (Ramkrishna, 2002; Randolph and Larson, 1984). Modeling and numerical solution of population balance models accounting for crystal shape has been discussed in the literature as well, e.g. Borchert et al. (2009a,b), Briesen (2006) and Ma et al. (2002, 2008).

The observation of crystal shapes using video microscopy has been investigated by different groups, e.g. Eggers (2008), Glicksman et al. (1994), Kempkes et al. (2008, 2010b), Larsen and Rawlings (2009), Li et al. (2006), and Patience and Rawlings (2001). The improvement of the quality of the measurements can in principle be done via the enhancement of the image quality or further development of the postprocessing algorithms. That is, either the hardware of the sensor is improved or extended or processing algorithms for image analysis are equipped with advanced techniques in order to apply it to data acquired from the (commercially) available equipment.

In principle, the evolution of crystal shape distributions can be observed in experiments and simulations of process models can be performed. Though the coupling of observation and simulation opens a wide field of applications in crystallization, e.g. optimiza-tion of control strategies and estimation of kinetic parameters

---

* Corresponding author. Tel.: +1 391 6110 350; fax: +49 391 6110 523.
  E-mail address: sundmacher@mpi-magdeburg.mpg.de (K. Sundmacher).

with regard to shape and size, only few work in this direction has been published, e.g. Patience and Rawlings (2001) and Wang et al. (2008). Given the fact that microscopic images can be acquired, the questions this paper aims to answer are: How can crystal projections easily be related to the crystal's 3D body in order to measure the shape distribution of a population in terms as it is used in the model? And how reliable is this information?

In the following, highly idealized dynamic models for single crystals and crystal populations are sketched (Section 2). The computed evolving distribution is artificially sampled and measured using an observer whose concept is worked out in Section 3. In Section 4 of this paper an evolving crystal population is tracked over the course of a whole simulated crystallization experiment and growth kinetics which have been used in the simulation are reestimated from the measurements. Section 5 concludes the paper in brief.

2. Process model

The form of a convex crystal can be described by the orientation $\mathbf{n}_i$ of its faces and their distances h_i to the crystal center, see Fig. 1. Since the face orientation is fixed (Miller indices) we recognize the vector $\mathbf{h} = (h_1, h_2, \ldots, h_n)^T$ as the geometrical state being a point in the geometrical state space $O \subset \mathbb{R}^n$. Presuming that symmetry-related faces have the same distance to the crystal center, the dimensionality of the state vector can be reduced considerably. For example, if the planes of the crystal as shown in Fig. 1 are allowed to assume an individual distance to the center each, a 26D state space is spanned. Instead, we deem all symmetry-related faces to have the same distance to the crystal center which is a model simplification but often justified in growth-dominated systems. This reduces the number of dimensions for the example of the crystal in Fig. 1 to three. We show the applicability of the presented methods throughout the paper using the example of a symmetric cubic crystal (cubic point group O_h) with faces {1 0 0} and {111}, which is a representative and demanding example system in the sense that the morphological variations are rich and many interfacial angles are equal between different pairs of faces.

In the remainder of this section the population balance is presented, Section 2.1. This population is grown in a batch crystallizer by natural cooling. For this case the temperature profile and the mass balance which takes the shape-dependent mass uptake into account is presented in Section 2.2. The population balance is discretized and can then be solved using simple ODE-solvers as described in Section 2.3.

2.1. Population balance equation accounting for shape evolution

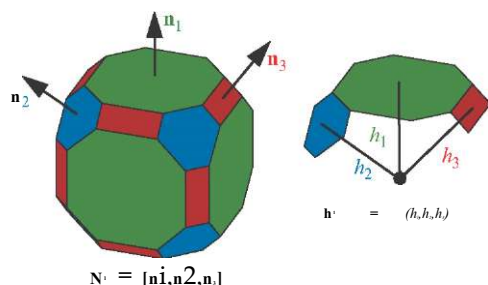Shape evolution of one crystal under the constraint of a constant number of faces within the domain $\mathbf{O}$ can be described by the vector differential equation (Borchert et al., 2009a; Taylor et al., 1992; Winn and Doherty, 2000):

$$\frac{d\mathbf{h}}{dt} = \mathbf{G}, \tag{1}$$

where $\mathbf{G} = (G_1, G_2, \ldots)^T$ is the vector of the perpendicular growth rates of crystal faces. If not only the shape of one crystal but also the shape distribution of a whole population of crystals has to be described by a model, the preceding equation does not suffice. This is at best done by using a population balance equation with which it is in principle possible to model various effects, for instance nucleation of new crystals, growth and dissolution, breakage and aggregation. In the following we assume that no nucleation, dissolution, breakage or aggregation takes place in the crystallization system. That is, only growth occurs and thus a highly idealized case is considered. This is also true when the observation scheme in a later part of this paper is tested: if this will not work for such an ideal case it is unlikely to work for more realistic cases.

Using the assumptions made above, the following multidimensional population balance accounting for crystal shape in a seeded batch crystallizer reads

$$\frac{\partial n}{\partial t} + \nabla_\mathbf{h} \cdot (\mathbf{G} n) = 0, \tag{2a}$$

subject to the initial, regularity and boundary conditions:

I.C. : $n(t=0, \mathbf{h}) = n_{seed}(\mathbf{h})$, \hfill (2b)

R.C. : $n(t, \mathbf{h}) \to 0, \quad |\mathbf{h}| \to \infty$, \hfill (2c)

B.C. : $\mathbf{n} \cdot \mathbf{G} n = 0, \quad \mathbf{h} \in \partial O$. \hfill (2d)

Of course, interesting dynamic solutions can only be produced if the seed distribution is different from the steady state solution to Eq. (2a). The main barrier for the application of such detailed population balance models is that growth and nucleation kinetics are unknown. A simple empirical growth model is given by

$$G_i = k_{gi} a^{g_i}, \tag{3}$$

where a is the supersaturation (Mullin, 2001; Myerson, 2002). The parameters k_{gi} and g_i are functions of solute composition and in practice not fully known. They are highly sensitive to certain process characteristics, in particular impurities and flow conditions. We shall be concerned with the determination of growth kinetics usable in population balance models. For this, observed distribution data as presented later in this paper are required.

Due to the occurring nonlinearities in the growth model, coupling to continuous phase balances and possibly nonlinear initial conditions, the population balance must be solved by numerical methods in most cases. The technique employed is presented after a short introduction to the used continuous phase balances.

2.2. Continuous phase balance equations

The mass balance of a batch system for the solute of which the crystalline phase is being assembled reads

$$\frac{dm_s}{dt} = -\rho_{cryst} \int n (\mathbf{G} \cdot \nabla V_{cryst}(\mathbf{h})) \, dV, \quad m_{solute}(t=0) = m_{solute,0}, \tag{4}$$

where ρ_{cryst} is the crystal density. $V_{cryst}(\mathbf{h})$ stands for the volume of the crystal as a function of the geometrical state which is calculated as described in Borchert et al. (2009a). The solute concentration is for practical reasons defined to be the mass of the
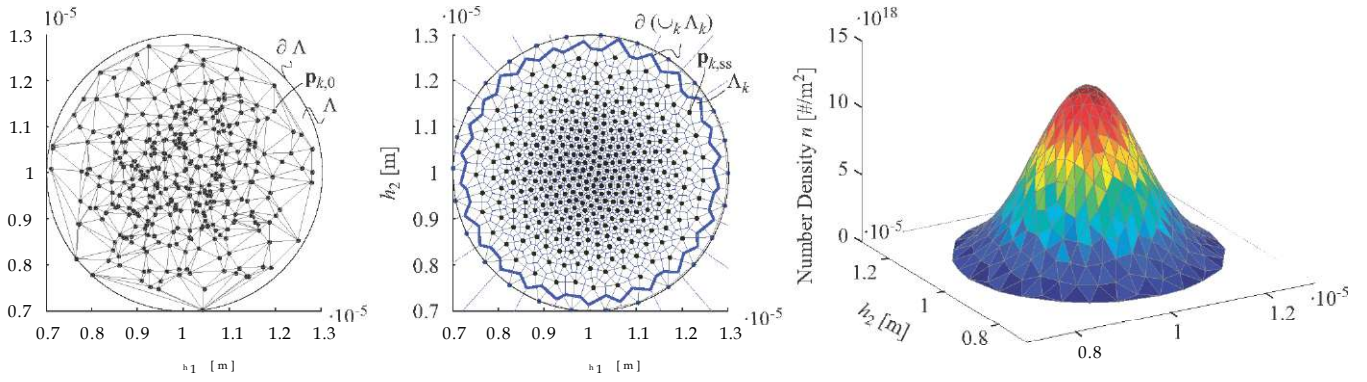

Fig. 1. A convex crystal is described by the orientation of its faces, $\mathbf{n}_i$, and their distance to the center, h_i.

$$\mathbf{h}_i = (h_1, h_2, h_3)$$

$$\mathbf{N}_i = [n_1, n_2, n_3]$$

Fig. 2. Meshing of the population region. Left: randomly distributed nodes. Middle: equilibrated mesh. Right: initial distribution on the unstructured mesh.

current amount of solute to the solvent mass (Mullin, 2001):

$$C = \frac{\text{-solute}}{\text{-solvent}}, \qquad (5)$$

where the equilibrium concentration as a function of temperature is an empirical relationship, e.g. a second order polynomial

$$C_{sat} = p2T^2 + p_1 T + p_0. \qquad (6)$$

The driving force for growth is the supersaturation which is defined here as the difference between the current and the equilibrium concentrations:

$$s = C\text{-}C_{sat}(T). \qquad (7)$$

From the energy balance the differential equation for the temperature is obtained as

$$V r c_p \frac{d}{dt} = f c_A O_j - T) + h c \text{ ffgr}. \qquad (8)$$

which is solved by

$$T(t) = T_J + (T_0 - T_J) exp(-kA/(V p c_p)t) \qquad (9)$$

for the case of natural cooling, $T_J = const.$, and negligible heat of crystallization. The parameter $kA/(Vpc_p)$ is set to a value reflecting the dynamics of our laboratory crystallizer.

2.3. Discretized population balance

In order to solve Eq. (2) simultaneously with Eqs. (4) and (9) we approximate Eq. (2) by a system of ordinary differential equations. A characteristic curve of (2) is a solution to the following ODE

$$\frac{d\mathbf{h}}{dt} = \mathbf{G}, \qquad \mathbf{h}(t = 0) = \mathbf{h}_0, \qquad (10a)$$

$$\frac{dn}{dt} = 0, \qquad n(t = 0) = n_{seed}(h0), \qquad (10b)$$

where $\mathbf{h}0$ is the starting point of the characteristic curve. In principle, infinitely many characteristic curves exist originating at points $\mathbf{h}_0 \in O$ which all together represent the exact solution to the population balance. In order to numerically approximate the solution, only n characteristic curves starting at selected points, denoted by $\mathbf{h}_{k,0}$, are considered:

$$\wedge = \mathbf{G}, \qquad \mathbf{h}k(t = 0) = \mathbf{h}k,0, \qquad (11a)$$

$$\wedge = 0, \qquad n_k(t = 0) = n_{seed}(\mathbf{h}0,k), \quad k = 1, \ldots, n. \qquad (11b)$$

The representative nodes $\mathbf{h}_{k,0}$ must be chosen in a way that they properly support the approximation of the true density. The

projection of n_{seed} on a regular, rectilinear or structured grid would allow a good numerical approximation. However, the number of grid points would be relatively large while the possibilities for local grid refinement are limited. Therefore, we used an unstructured mesh which was generated with an algorithm along the work of Persson and Strang (2004).

Firstly, the geometry of the region of interest must be represented. The choice of this region, denoted by $\mathbf{A} \subset \mathbf{O}$, is in our case determined by the number density function, i.e. only regions with sufficiently high number densities are taken into account. Preliminary nodes $\mathbf{p}_{k,0}$, $k = 1, \ldots, V$ are distributed within the region and passed to a Delaunay triangulation routine which connects the nodes, see Fig. 2 (left). This formation is now considered to be a truss with nodes and elastic bars which are the connecting lines between nodes. Due to an interaction between connected nodes they are moved according to the ODE:

$$\mathbf{J}!\mathbf{P} = \mathbf{F}(\mathbf{p}), \qquad \mathbf{p}(t = 0) = \mathbf{p}0, \qquad (12)$$

with $\mathbf{p} = [\mathbf{p}_1, \ldots, \mathbf{p}_v]$ and $\mathbf{p}_0 = [\mathbf{p}_{1,0}, \ldots, \mathbf{p}_{v,0}]$. The velocity function $\mathbf{F}$ allows to manipulate the mesh with regard to spacing between nodes. We have chosen this function such that the equilibrium length between connected nodes is small in parts of the state space with high number densities and sparsely covered with nodes in regions with smaller densities. Note that t and $\mathbf{F}$ have no physical meaning, though a physical analogy has been used for illustration.

The steady state solution of (12), $\mathbf{F}(\mathbf{p}_{ss}) = \mathbf{0}$, provides well distributed points $\mathbf{p}_{k,ss}$. A Voronoi tesselation is taken out assigning a cell L_k to the node $\mathbf{p}_{k,ss}$, see Fig. 2 (middle). All nodes with a cell of finite size are taken as starting points of characteristic curves. If the mesh parameters are well chosen, this excludes nodes which were moved towards the boundary ∂A. I.e. the remaining network of n < V nodes covers a region in the interior of $\mathbf{A}$, see Fig. 2 (middle).

We have now generated a well distributed set of nodes and are thus ready to write down the discretized initial distribution:

$$_{k,0} - _{k,ss} \qquad (13a)$$

$$n_{seed}(\mathbf{h}0,k) = \frac{1}{\Delta V_{A,k}} \int_{A_k} n_{seed}(\mathbf{h}) \, dV_h, \qquad k = 1 \ldots V \qquad (13b)$$

with the cell volume

$$DV_{A,k} = \int_A dV_{\mathbf{h}}. \qquad (14)$$

The discretized seed distribution used for the simulation presented in the subsequent section is shown in Fig. 2 (right).

2.4. Numerical simulation

The discretized population balance equation, (11), and the mass balance (4) are discretized in time domain using the forward Euler scheme:

$$-solute, j+1 \quad - -solute, j \cdot P_{cryst} \sum_{k-1}^{v} n_{u}\{G \cdot W_{cryst}(\mathbf{h}kj))AVL, kAt, \quad (15a)$$

$$\mathbf{h}k, j+1 \quad - \mathbf{h}kj + G(Sj.1) At. \quad (15b)$$

An alternative to this simple discretization in time could be a higher order Runge-Kutta schemes (implicit or explicit). The temperature profile as given in Eq. (9) (natural cooling) has been used. This system is easily implementable (e.g. in MATLAB) and simulation studies can be taken out of which one is depicted in Fig. 3. The seed distribution in the bottom left corner of the state space is moved along a nonlinear curve which is the result of the coupling to the mass and energy balance. Initially, the system is just saturated and then cooled approaching ambient temperature and thus inducing supersaturation. This drives crystal growth
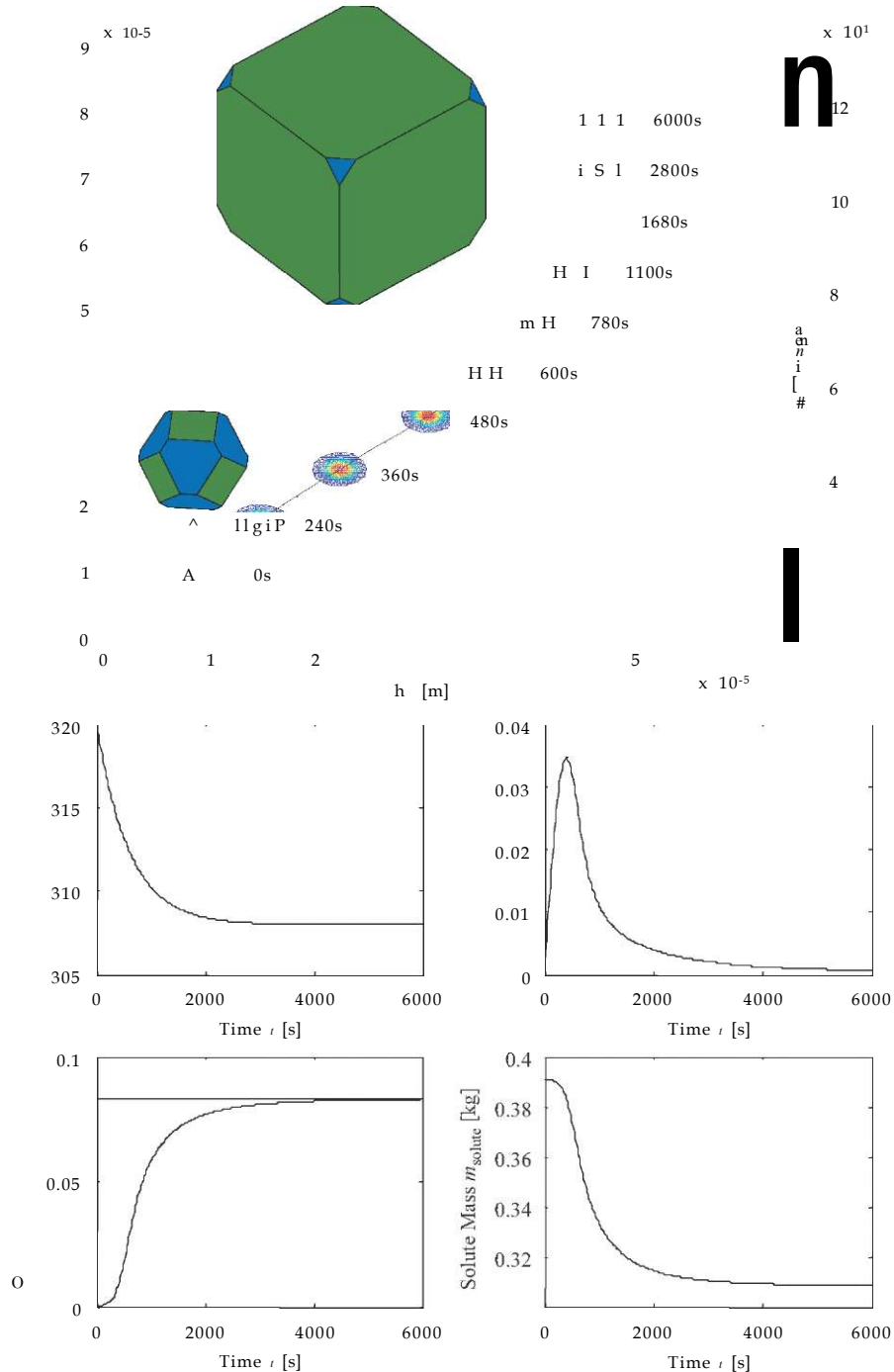

Fig. 3. Simulation results for a seeded batch crystallizer. Left: evolution of the number distribution with representative crystal shapes of the seed population and the final population. Right: continuous state variables.

Table 1
Parameters of the simulation.

h_{seed}	Mean state of seeds (m)	$(10^{-5}, 10^{-5})^T$
s_{seed}	Standard deviation of seeds (m)	$diag(10^{-12}, 10^{-12})^T$
	Number of crystals (#)	8.29×10^5
cc	Initial concentration (kgsolute/kgsolvent)	$c_{sat}(T0)$
$T0$	Initial temperature (K)	320
T	Jacket temperature (K)	308
$m_{solvent}$	Mass of solvent (kg)	1
$(k_{g,1}, k_{g,2})$	Growth parameter (m/s)	$(6 \times 10^{-5}, 2 \times 10^{-6})$
$(g1, g2)$	Growth parameter	(2,1)
ρ_{cryst}	Density of crystal (kg/m^3)	1160
$kA/(V \rho_c c_p)$	Crystallizer parameters (1/s)	1/600
$(p0, p1, p2)$	Solubility parameter	$(3.68, -2.82 \times 10^{-2}, 5.58 \times 10^{-5})$

leading to a transfer of dissolved material to the crystalline phase. At time $t = 2000$ s the main crystallization phase is completed, though there is still some degree of supersaturation left even at the end of the simulation at 6000 s. Parameters which were used in this case study are specified in Table 1.

3. Crystal observation

Modeling of a crystallization process like in the previous section is only useful when it can be equipped with proper kinetic expressions. Due to the complexity of numerous interacting process variables governing the dynamics of crystallizers, the quantitative knowledge of kinetic parameters as a function of a full set of process parameters is usually not available. Particularly crystal shape can be perturbed dramatically by minor changes in impurity concentration or simply through the switching to a different supersaturation level (Boerrigter et al., 2004; Gilmer, 1980; Weissbuch et al., 1991, 1995). Because of the high sensitivity it is necessary to observe a crystallization process with regard to the features of interest if they are critical. However, compared to size analysis the monitoring of particle shape is far less developed for good reasons of which in our opinion the most important ones are: (i) Sensors enabling the quantification of the full 3D shape, examples are given below, require thorough sampling preparation and are costly. (ii) In situ and ex situ microscopes deliver a 2D information from which it is rather difficult to extract quantitative information.

The precise, non-parametric reconstruction of the 3D crystal shape can for instance be obtained from variants of transmission electron microscopes (Koster et al., 2000; Weyland et al., 2001), by processing of image stacks from optical microscopes (Castro et al., 2003) or tomographic methods, e.g. Jerram et al. (2009) and RJL Micro & Analytic GmbH. Especially tomography offers promising applications with regard to shape characterization and quantification of dispersed phase systems in general. However, the application of tomography is costly and specialized staff is required to ensure efficient operation.

Though the full 3D shape information is desirable to be available in future devices we focus on extracting useful information from a single 2D projection of the particle. The main advantage is the relatively simple probe operation and handling. Classical optical microscopes offer the least costly method to acquire crystal images. But for this, a careful sample preparation is necessary which may alter the crystals. Further, for direct process control it is desirable to record and process crystal images online or even inline. Commercially available in situ sensors, e.g. PVM from Mettler Toledo (2001) or ParticleEye from Hitec Zang, can be installed directly in standard laboratory crystallizers. Ex situ sensors on the market, for instance from Sympatec (Qicpic) Sympatec GmbH or Retsch (Camsizer) Retsch Technology, have

usually a better image quality than in situ probes which comes at the cost of an additional sampling loop. The group of Mazzotti developed an ex situ sensor which was upgraded to acquire crystal images from two different perspectives, enabling the access to a wealth of information which makes it easier to relate to the real geometry because two different projections of the same situation are available (Kempkes et al., 2010a).

When the 3D geometrical state of the crystal has to be reconstructed from a single 2D image, only model-based methods can be used in contrast to non-parametric shape quantifications of true 3D sensors. The model is needed because features of the projection can then be used as an indicator for the 3D shape, that is, the space of possible objects throwing the projection is confined to those which are as well obtainable by the shape model. We use for this the crystal geometry model in terms of the state vector **h**. Clearly, **h** is the quantity we wish to measure. An additional difficulty to this is that the orientation of the crystal in space has a decisive influence on the shape of the photoprojection but is, depending on the sensor, a more or less stochastic process. In practice, the acquisition of images is performed with different probes as discussed above. For the design and evaluation of the estimation scheme, however, synthetic images have been generated so that the measured state and orientation can be directly and quantitatively compared to the actual state.

The matching of a crystal projection with the actual crystal shape constitutes a highly nonlinear optimization problem which is tractable by different means. Using continuous, gradient-based optimization techniques in orientation and geometrical state space would be a possibility, but the starting point must be chosen well beneath the true values in order to avoid the solver to run into a local minimum. Global optimizers using heuristic or stochastic techniques require a large number of relatively costly function evaluations and thus the estimation of a large number of crystal shapes becomes infeasible. Therefore, we have used a lookup table in which the shape descriptors as functions of state and orientation are precomputed. With this it is also possible to go the inverse way: the determination of state and orientation which corresponds to a specific descriptor. This is a commonly used approach in object recognition (DeMenthon and Davis, 1992) and has for instance been used to distinguish crystal polymorphs (Calderon De Anda et al., 2005; Li et al., 2006).

The rest of this section is organized such that at first descriptors for projections of crystal shapes are introduced and how they are extracted from an image. A simple state estimator is sketched and applied to measure a shape distribution from computer-generated crystal images.

3.1. Crystal projection and descriptors

Crystals which are observed with microscopes are viewed from one perspective and essentially we see a projection of the

particle. Though there is a huge information content in the grayscale landscape within the particle, this information is particularly sensitive to the system under consideration, complex to analyze and difficult to interpret by quantitative means. Therefore, we focus on the shape of the projection which is given by its boundary curve. This curve is determined by the shape of the crystal and its orientation in space, see Fig. 4.

For a convex crystal the boundary curve is determined by the convex hull of the projection of the crystal vertices which are connected by straight lines. That is, it is fully determined by (i) the length of straight lines, d, connecting projected vertices and (ii) angles, j_i, between them, see Fig. 6. This is the basis for a set of shape descriptors. In order to obtain size-independent descriptors, the dj's are scaled that the sum of the resulting boundary perimeter is one:

$$dj = \frac{}{\sum_{f=1} dj} \qquad \sum_{j=1}^{E} d = 1. \qquad (16)$$

where n_o is the number of straight lines of the boundary. Of course, n_o is a function of shape and orientation and not fixed for a particular morphology but can be determined by an algorithm, e.g. Cole (1966). The quantity

$$s = d \qquad (17)$$

is the scaling factor relating original with scaled measures. Rotation invariance is accomplished by ordering the dj's in descending sequence:

$$d_1 > d_2 > \cdots > d_n. \qquad (18)$$

For the special case $dj = dj_{-1}$ the ordering is chosen such that dj_{-1} follows dj in clockwise direction. From the angle between two boundary lines, $j = +(dj.dj_{-1})$, $j_n = +(d_n.d)$, the absolute value of the cosine is taken and together with the scaled boundary lines assembled to a descriptor vector:

$$\mathbf{d} = \mathbf{f}_{descr}(\mathbf{h},W) = (d1 .|\cos j1| \qquad dn. |\cos j \; 1)^\top \qquad (19)$$

It is obvious that a loss of information is caused by the projection and in general there is not a unique mapping from the descriptor vector $\mathbf{d}$ to the state vector $\mathbf{h}$. I.e. the same set of descriptors can be produced from different states and orientations so that the inverse

$$(\mathbf{h},W) = \mathbf{f}_{descr}(\mathbf{d}) \qquad (20)$$

must not necessarily be unique. However, when the geometrical state is given, it is an easy task to project the vertices on a plane and determine the descriptor vector $\mathbf{d}$. An overview of how to obtain descriptors from state and orientation is depicted in Fig. 5.

The main idea of the next part is to build up a lookup table or database of descriptors for which the geometrical state is known. By comparing a measured descriptor vector to this data set, a guess can be made in which crystal shape is at hand.

3.2. State estimation

The lookup table (lut) is compiled from numerical experiments: The descriptor vectors $\mathbf{d}_{lut,i}$ are acquired for boundary curves of crystals with randomly chosen states $\mathbf{h}_{lut,i}$ and orientations $W_{lut,i}$. The whole set of descriptors and states

$$D = \{\mathbf{d}lut,1 \qquad \mathbf{d}lut.n_{lut}\}, \qquad (21a)$$

$$= (\,lut,1. \; \cdots \; .lut.n_{lut}). \qquad (21b)$$
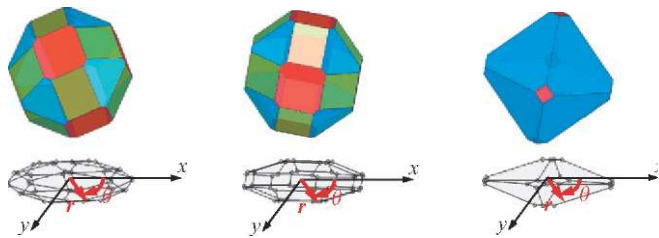
$$= (\,lut.1 \qquad lut.n_{lut}) \qquad (21c)$$

Fig. 4. Crystal projections: the same crystal shape projected from different perspectives (left and middle). Different shapes photographed from the same direction (middle and right).

$$\mathbf{d} = (o_y\cos \langle jo \rangle \; {}^{\wedge}2 {}^{\wedge}i^{\cos} \; {}^{\wedge}2! \; {}^{\wedge}3 {}^{\wedge}1^{\cos} \; \text{CR}, \; {}^{\wedge}4, |\cos \; \text{P}4|, \; d_n,|\cos \qquad d_6,|\cos$$
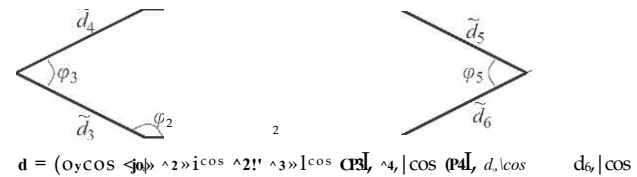
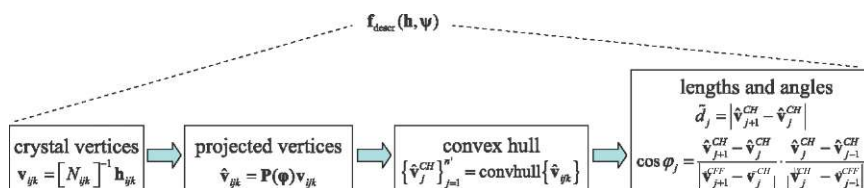Fig. 6. Descriptors of the crystal projection boundary.

Fig. 5. Relationship between crystal boundary and its descriptors.

is the boundary descriptor lookup table, see Fig. 5. A state estimator uses this database and compares a measured descriptor **d** to those given in the table:

$$e(\mathbf{d},\mathbf{d}_{lut,j}) = dist(\mathbf{d}-\mathbf{d}_{lut,j}), \quad \mathbf{d} \wedge - e\ D,\ j = 1 \qquad n_{lut}, \quad (22)$$

where the distance measure is of a special form since the number of descriptors can vary, depending on how many straight lines are exposed on the boundary of the descriptor from the database or the measurement. For two vectors **x,y** the first entries are compared:

$$dist(\mathbf{x}-\mathbf{y}) = \frac{1}{J2} \sum_{j=1}^{n-1} |x_j - y_j| \mathbf{I}, \quad \mathbf{x}\ e\ R^n,\ \mathbf{y}\ e\ R^m\ n < m \quad (23)$$

The sum goes from 1 to *n*-1 (and not to *n,* the number of components of the shorter vector **x)** because the last entry of the descriptor vector, Eq. (19), contains for **x** the cosine of an angle between the shortest ((n-1) th entry) and the longest (1st entry) line segment of the projection boundary. On the other hand if the descriptor vector **y** is larger than **x,** *m > n,* the nth entry is a cosine between two consecutive (in terms of length) line segments. That is, the two entries refer to different quantities which should not be compared. Furthermore, the reason for taking descriptor vectors of the database into account which have a different dimension than the measured one is that only slight changes in the orientation can cause a switch in the descriptor dimension because the number of edges projected onto the camera chip can change by a small orientation perturbation. However, the additionally appearing features are rather small in magnitude and in case that no descriptor vector has been inserted which was obtained exactly from that orientation, the closest entry in the database can still be found even though a qualitative jump lies between database entry and measured descriptor.

Finally, the entry in the database which deviates least from the measured one is taken as the hit and the geometry of the crystal can be identified:

$$\mathbf{h}_{est,sc} = \mathbf{h}_{lut,j} \quad |\ \mathbf{h}_{lut,j}O < \ |\ \mathbf{h}_{lut,k}\ \quad \mathbf{h}_{lut,k} \cdot D, \quad (24)$$

where **h**est,sc must be rescaled to original dimensions using the scaling factor, Eq. (17):

$$\mathbf{h}_{est} = \mathbf{f}_{est}(\mathbf{d}) = \mathbf{s}\mathbf{h}_e. \quad (25)$$

We shall now be concerned with the accuracy of this estimation scheme for the ideal case that the descriptors can be calculated perfectly.
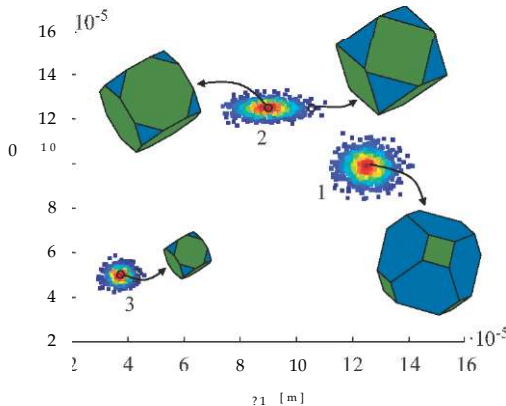
3.3. Performance of the estimator

A population consisting of three subpopulations is examined which are all normally distributed with means and variances

$$= (1.25,1)\ x\ 10^{-4}, \quad s_{1,11} = 5\ x\ 10^{-6}, \quad s_{1,22} = 5\ x\ 10^{-6},$$

$$= (0.9,1.25)\ x\ 10^{-4} \quad s_{2,11} = 1.25\ x\ 10^{-5} \quad s_{2,22} = 2.5\ x\ 10$$

$$l3 = (0.375,0.5)\ x\ 10^{-4}, \quad ff_{3,11} = 2.5\ x\ 10^{-6} \quad s_{3,22} = 2.5\ x\ 10 \quad (26)$$

and vanishing covariances. From each subpopulation 1000 samples are taken, see Fig. 7 (left), randomly oriented and the descriptor vector $\mathbf{d} = \mathbf{f}_{...,\mathbf{r}}(\mathbf{h},\mathbf{W})$ is determined. The procedure described in Section 3.2 is used to reestimate the geometrical state. As depicted in Fig. 7 (right) the features of the population can be reconstructed well. Especially the distinction between different subpopulation is possible in a reliable way. The error of the estimate of the *j*th particle with respect to the quantity h_k is measured by

$$_{jk,est} - \frac{_{k,est} - _{k,true}}{_{k,true}} \quad (27)$$

Fig. 8 shows the error distribution for each of the three populations for both quantities *h*1 and *h*2. It can be seen that the majority of the estimates comes with an error below 2%. The standard deviation of the error distribution ranges from 2.41 x 10^{-3} to 1.15 x 10^{-2}. The error distribution indicates that the estimation scheme is not working equally well in all areas of the state space. Particularly population 1 can be identified with a better quality than populations 2 and 3. The accuracy of the estimates of *h*2 in populations 2 and 3 is rather poor compared to estimates of *h*1. This is due to the lower prominence of the *h*2-faces on the crystal surface, see examples of these shapes in Fig. 7 (left). Therefore, features of *h*2-faces are less expressed in the projections and thus the distinction between subtle differences of *h*2-features requires highly resolved lookup tables to achieve better results. However, we find that a standard deviation of the error, denoted by s_e, of less than 3% is sufficient for our purposes. This result has been achieved with a lookup table which was made up of $n_s = 1000$ different state vectors whose resulting crystal shapes are each photographed from $n_W = 300$ different random directions.

Assessing the improvement of the estimation quality with varying configurations of the lookup table can be done by considering the standard deviation of the error distribution s_e. This is depicted for population 1 in Fig. 9 with a variation of the
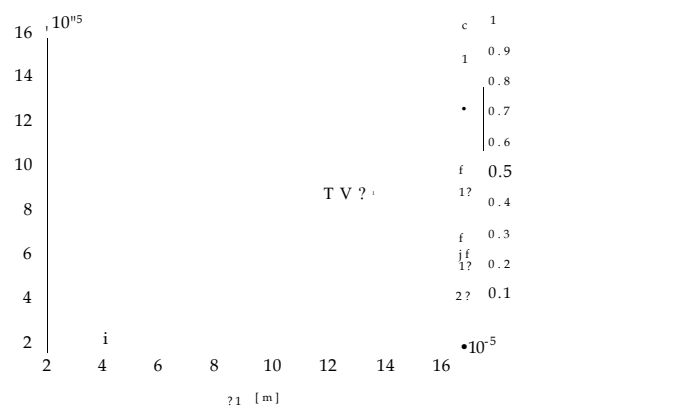

Fig. 7. Estimation of a trimodal population consisting of 1000 crystals for each of the three populations. Left: scatter plot of samples from the original population including graphical representations of selected crystal morphologies. Right: scatter plot of the estimated population.
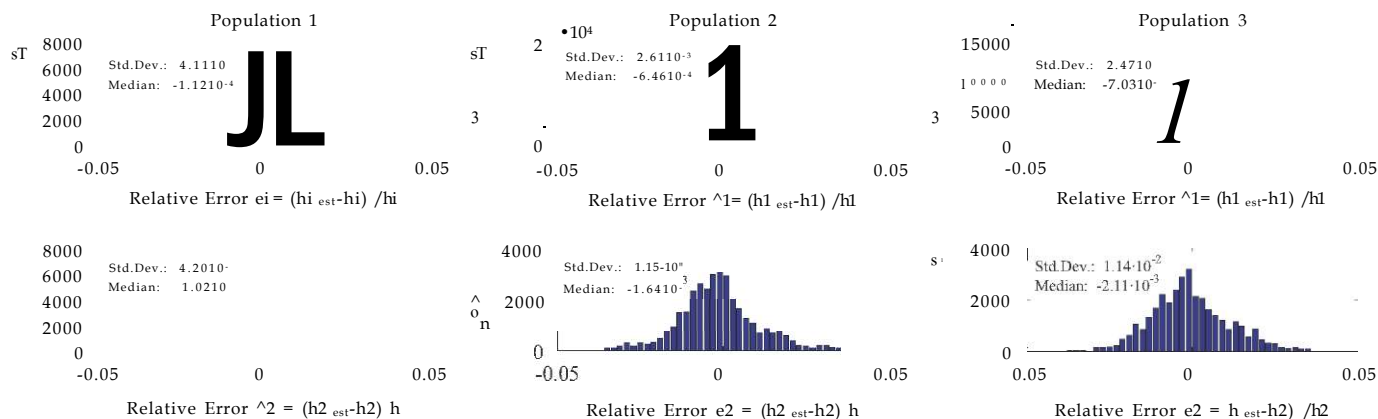
Fig. 8. Error distribution based on the comparison between real and estimated state of individual crystals for the three subpopulations shown in Fig. 7
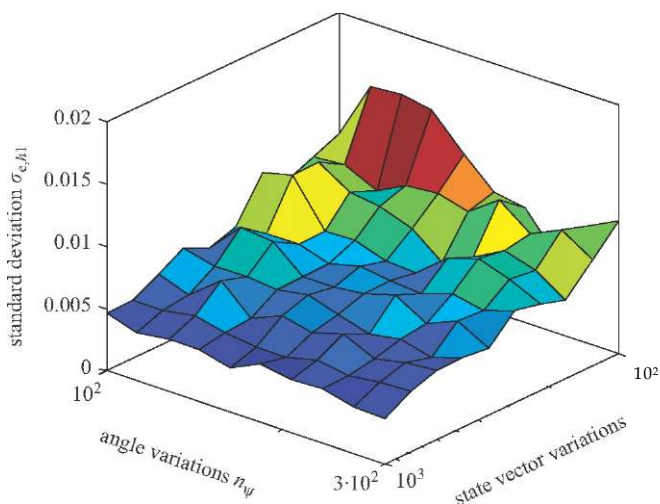

Fig. 9. Error between given and estimated shape distribution as a function of database size.

lookup table size from $n_h = 100\ldots 1000$ and $n_w = 100\ldots 300$. As expected, a blow-up of the table size improves the accuracy of the estimates. Particularly an increase in the number of states, nh, reproduces the population in a better way while the influence of the number of orientation variations nw is weak. It is clear now that the estimation scheme as described in Section 3.2 works in principle if the shape descriptors **d** are given. Now, it is an additional task to find a procedure which determines the shape descriptors from pixel and thus possibly grainy images. This topic is addressed next.

3.4. Descriptors from pixel images using the Hough transform

On a pixelized image the boundary is given by the coordinates of n points. We are looking for pieces of the boundary which lie on straight lines. In principle, a line can be drawn through every pair of points and subsequently all subsets of points that are close to particular lines can be determined. This approach would involve the detection of n^2 lines and the computation of n^3 comparisons which is computationally expensive. The Hough transform instead offers a more favorable method at a computational cost increasing linearly with n. The illustration of the Hough transform follows mainly the textbook of Gonzalez (Gonzalez and Woods, 2008).

The equation of a straight line is

$$r = x \cos y + y \sin y, \tag{28}$$

where r is the distance from the line to the origin and **y** is the angle of the shortest vector from the origin to a point on the line. Let (Xj, yj) be the coordinates of a pixel. Infinitely many lines pass through (Xj, yj) which all satisfy Eq. (28) for a set of parameters $(r, \mathbf{y})$ which is a curve in $(r, \mathbf{y})$-space:

$$r = Xj \cos y + yj \sin y. \tag{29}$$

Clearly, another equation for a point (x_i, y_i) would then allow to determine the intercept of the parameter curves giving parameters $(r', \mathbf{y'})$ of a line passing through both points, see Fig. 10. The parameter space is now subdivided into accumulator points, e.g. A_{lm} as depicted in Fig. 10 (right). Initially, the accumulator points are all set to zero before we run through all boundary pixels. Then, for every boundary pixel the curve (29) is evaluated on the accumulator grid. If the curve passes through the neighborhood of $(P_l, \mathbf{y}_m)$, A_{lm} is increased by one. At the end of this procedure A_{lm} points of the boundary lie on a straight line with parameters $(P_l, \mathbf{y}_m)$. Roughly speaking, the maxima of the accumulator points correspond to the most prominent straight lines in the original image, see Fig. 11.

From the lengths and orientations of the computed boundary section, descriptors **d** can be calculated and used to determine the shape of the photographed crystal. In Fig. 12 the previously used trimodal population has been projected on a plane with a resolution of $(10 \quad m, 10^{-6} m)$ per pixel. It can be seen that the main features of the population can be recognized from the measurements but with a lower quality (Fig. 13) as if the descriptors are obtained from infinitely resolved planes (Section 3.3). This is mainly due to imperfect line detection through the Hough transform which can be difficult to operate at different length scales simultaneously. That means, subtle features being important for small projections are undesired to be detected in larger projections. By improving the line detection algorithm we think that for sufficiently resolved projections the hit rate can still be improved, then comparable to the ideal case.

4. Connecting simulation and estimation

So far we have been concerned with the modeling of a crystallization process in Section 2 and observation of crystals in Section 3. The application of the developed observation techniques to samples taken from the evolving number distribution is conducted in this section. It serves to evaluate requirements which have to be met to observe the crystallization process
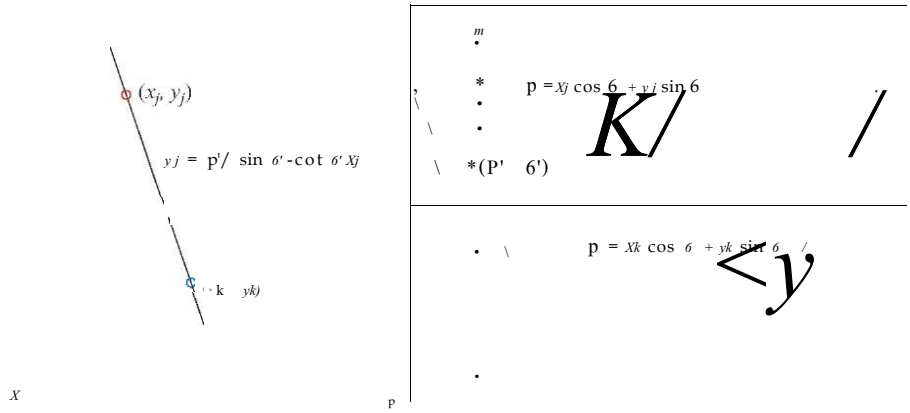
Fig. 10. Line passing through two points (left) and parameter curves of all possible lines passing through the respective points (right). The point of intersection, *(r',9'),* yields the parameters of the straight line passing through both points in the original (x,y)-space.
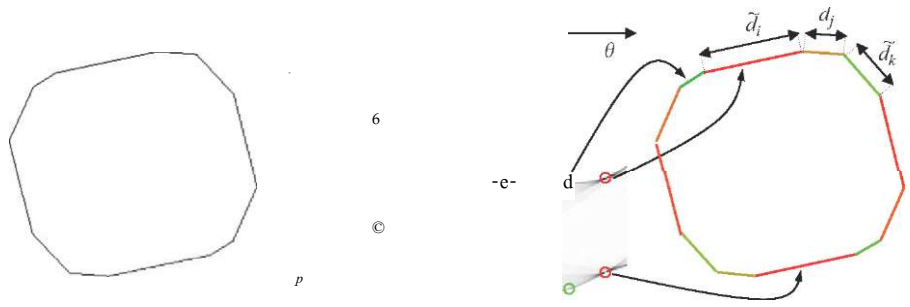

Fig. 11. Extraction of straight lines of a boundary curve. Left: crystal boundary from a simulated photograph. Middle: intensity image of the accumulator points in the parameter space. Right: identified straight lines.
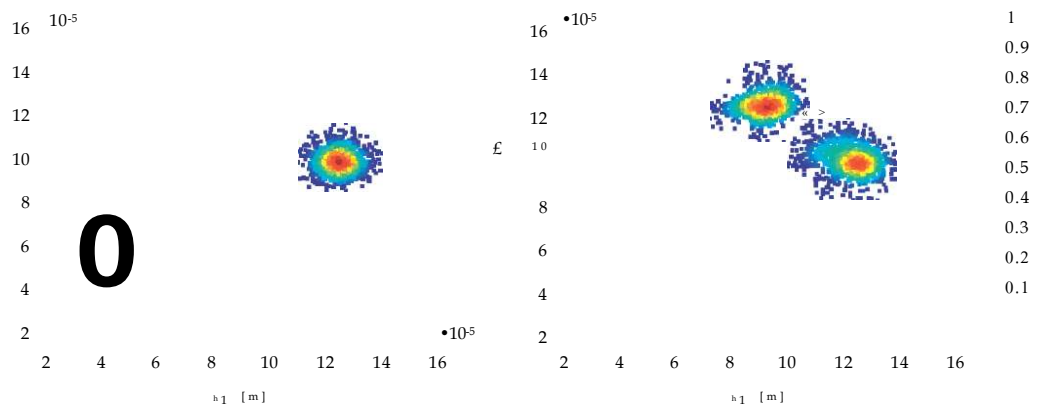

Fig. 12. Estimation of a trimodal population consisting of 3000 crystals. Left: scatter plot of original particle population. Right: scatter plot of the estimated population. Examples of crystal shapes in this region of the state space can be seen in Fig. 7.

adequately with regard to number of samples and frequency of sampling. Especially for the identification of kinetics used in population balance equations it is important to get an idea of how the measurement process must be dimensioned in order to be able to extract information of good quality.

The shape distribution evolution as computed in Section 2.4, shown in Fig. 3, has been sampled at ns — 10 instants, where each sample comprises nc — 200 crystals. Fig. 14 (left) depicts the sampled population densities. The density contour plots are computed with a histogram smoothing algorithm (Eilers and Goeman, 2004). Applying the state estimation scheme as discussed above with a lookup table with n_k — 1000,$n^$ — 300, the estimated populations are obtained as sketched in Fig. 14 (right).

Unsurprisingly, the main features of the evolving distribution can be recognized in the observed population.

The quantitative observation of a crystallization process is of value in its own right. But even more desirable is the extraction of kinetic data from an observed process. Since only crystal growth has been taken into account, we aim at restoring the growth rates which have been used in the simulation. Clearly, we could try to apply the estimation technique directly to real experiments and determine growth rates. However, we find it essential to test an observation and estimation scheme against artificial experiments to gain information on the reliability of the applied technique.

A first, model free determination of the growth rates is based on the evaluation of particle size change between two measurements.
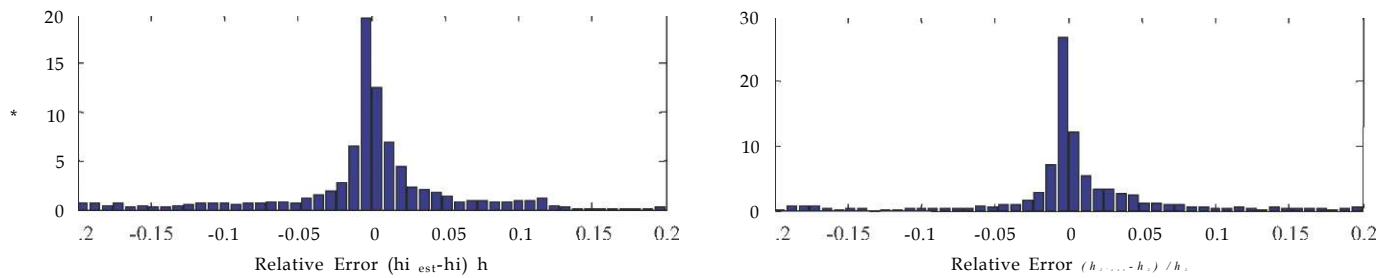
Fig. 13. Error distribution of estimation population photographed on pixelized plane.
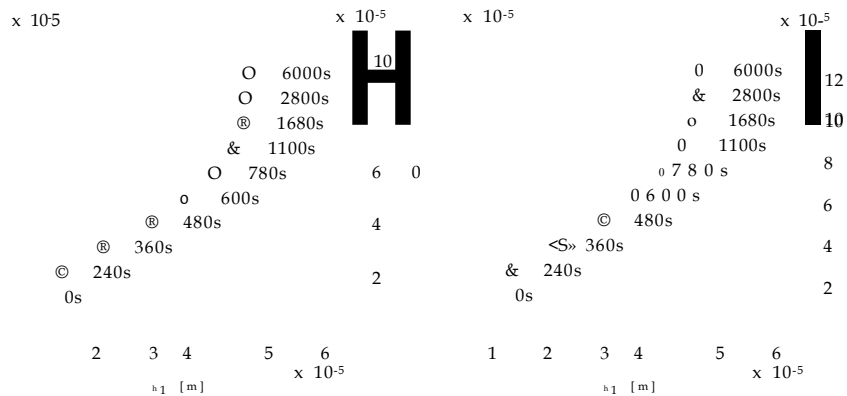

Fig. 14. Evolution of the number distribution as in the simulation study. Left: sampled distribution. Right: estimated distribution from photos with infinite resolution. The simulated evolution of the shape distribution is shown in Fig. 3 together with representative shapes at the beginning and end of the crystallization.
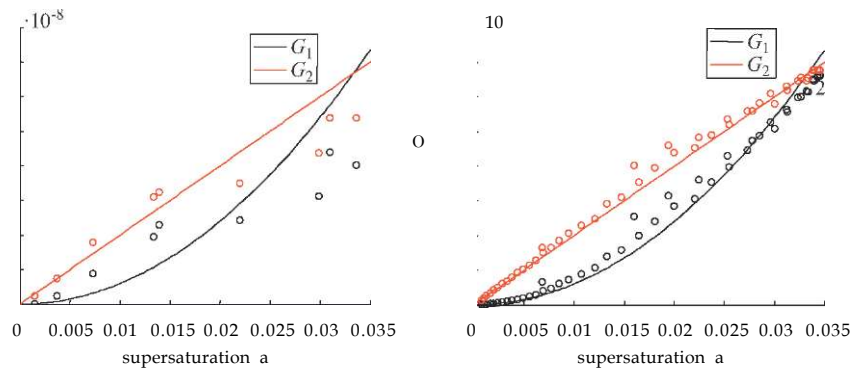

Fig. 15. Estimation of growth rates from simulated experiments, (-) true kinetics, **(o)** measured kinetics. Left: taking only 10 samples over the course of a whole crystallization experiment does not suffice to determine the growth kinetics. Right: with 50 samples the true kinetics can be approximated well.

By differentiating the mean particle size of the seed crystals

$$\overline{h}_j = \frac{1}{n} \sum_{k=1}^{n} V \quad {}_j = {}_{\cdots \cdots ,"\gg} \tag{30}$$

numerically with respect to time

$$G \quad \frac{h_{j+1} - h_{,-}}{ti+1 - ti} \tag{31}$$

and measuring the supersaturation or concentration with the same sampling rate we can plot the measured growth rate versus supersaturation and get an impression of the growth law. This is depicted in Fig. 15. If only samples at 10 instances are taken, it is not possible to estimate the underlying growth law properly.

According to our systematic study, using 50 samples enhances the quality considerably.

If a growth model of the form (3) is assumed, the measured evolution of the mean crystal size h_j can be compared to the mean crystal size following from the growth law together with supersaturation measurements:

$$\mathbf{h}_{\text{mod}}(t) = \mathbf{h}_c + \int \mathbf{k}_g s_i \, dt. \quad \mathbf{k}_g = (k_{g,1}, k_{g,2})^\top. \quad \mathbf{g} = (g_1, g_2)^\top. \tag{32}$$

The deviation between measured and modeled mean size evolution is quantified by the objectives

$$e_1 = \sum_{j=1}^{n} {}^{\wedge}Z(h_{1j} - h_{1,\text{mod}j})^2, \tag{33a}$$
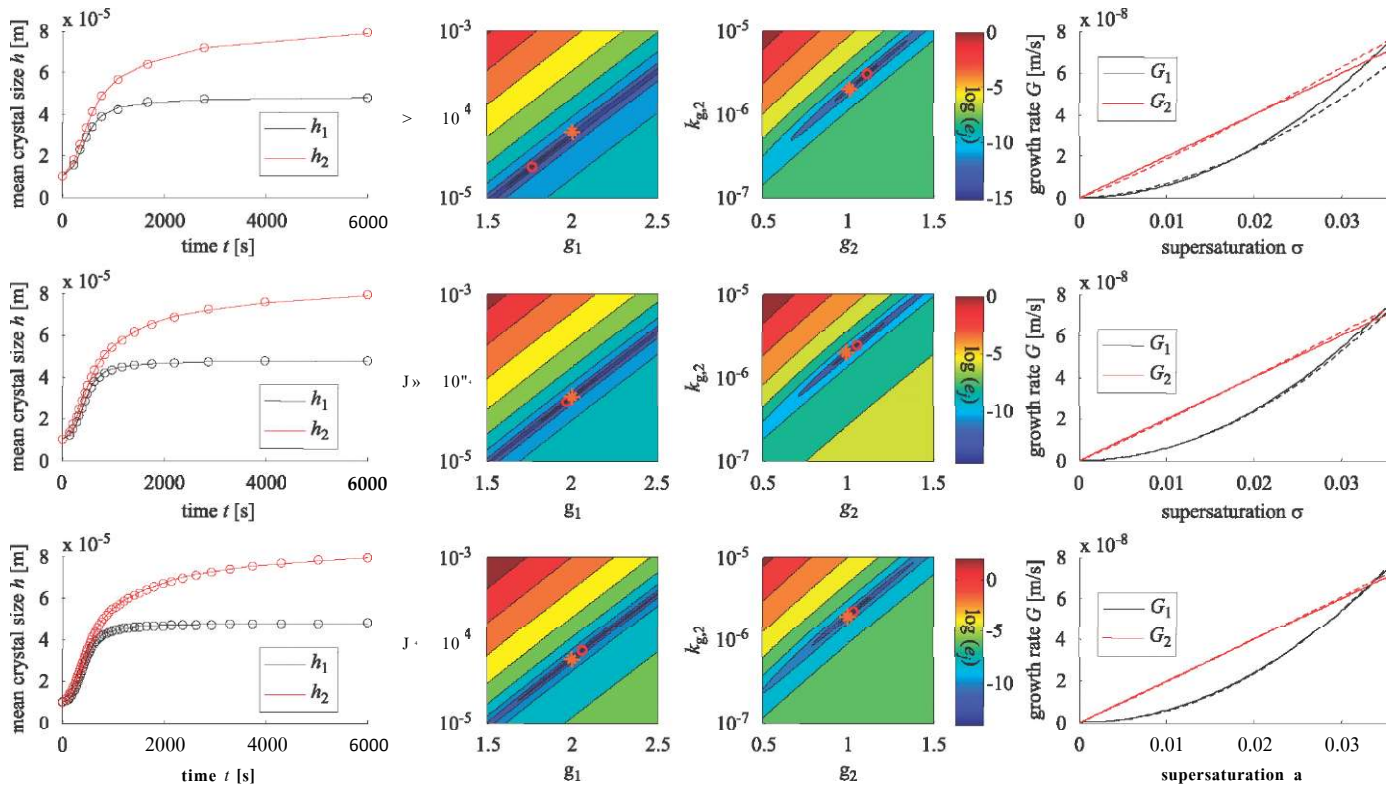
Fig. 16. Estimation of growth rates with underlying model using a varying number of samples. Upper row: 10 samples, middle row: 20 samples, bottom row: 50 samples. Left column: evolution of mean crystal size estimated from simulated experiments. Middle columns: contour plots of the objective e_j as a function of growth parameters. The true parameter used in the simulation is indicated by a star, whereas the minimum of the objective is marked by a red circle. Right column: true (-) and estimated (- -) growth laws. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

$$e_2 = \sum_{j=1}^{ns} X(\bar{h}_{2,j} - \bar{h}_{2,mod,j})^2. \tag{33b}$$

In Fig. 16 the growth laws which are estimated from the minimization of the objective functions are depicted. It can be seen that with an increasing number of samples the accuracy of the estimates increases. Compared to the direct numerical differentiation of the mean crystal size which has been used in Fig. 15, the model-based estimation yields much better estimates even with relatively few samples. The estimation based on 20 samples (middle row) almost perfectly matches the true kinetics. The objective shows, independent of the number of samples, stretched valleys which are undesirable when optimization algorithms are applied. Also, if it comes to the measurement of real data, the acquisition of images and measurement of shape distributions are further complicated by crystal shapes which are not as ideally formed as in the simulation. This can involve deviations from symmetry, formation of aggregates or broken crystals which all lead to more complex (that is, a wider variety of) crystal shapes whose formation are stochastic processes. This means for example, the aggregate geometry formed by the clustering of a number of crystals is due to its complexity essentially a random process. That is, the stochastic process of crystal orientation, which has been included in our analysis, is further superimposed by other stochastic processes. Hence, the data quality may further decrease and therefore stretched valleys of the objective in the parameter space further complicates the judgement over the quality of the estimated parameters. We believe that this can be remedied by the model-based analysis and redesign of the conduction of the crystallization experiment of which major building bricks have been fabricated here.

5. Conclusions

The achievements of this paper are twofold. At first we have proposed an efficient numerical solution technique based on the method of characteristics. A meshing algorithm has been employed for this which allows the discretization of the seed population so that the starting points of characteristic curves can be determined in a rational way. Explicit Euler discretization in time makes it very easy to implement the resulting equations and gives numerically stable results for nonstiff equations.

Secondly, an observation scheme for faceted crystals has been developed. It is based on the notion that the projections of crystals as recorded by microscopes (classical laboratory microscopes, in- or ex situ probes) are strongly related to the actual 3D shape but randomized by the stochastic process of orientation. Since the inversion of observed features of the projection to the actual shape is a complicated reconstruction task and often not possible we proposed to use a lookup table. In this table the descriptors of the projections are tabulated together with the geometrical state and orientation with which they were computed. A simple distance function evaluating the closeness to tabulated values is used to find a matching entry in the lookup table. This value is then taken as the estimated geometrical state. Robustness is guaranteed in the sense that - compared to more sophisticated optimization techniques - the algorithm requires a fixed number of operations per estimate and cannot run into convergence traps. It may sometimes not hit the true shape exactly. This comes partly from the finite resolution of the lookup table but is also due to ambiguous descriptors which may be produced from quite different shapes and orientations. But overall, the proposed technique has been shown to work well in different examples with morphologically mixed populations.

For the application to real systems two things must be addressed beforehand. The continuous feeding of new characteristic curves to the numerical scheme is required when nucleation plays a role in the system under consideration. Further, the state estimation scheme must prove to be robust against noisy boundaries in pixelized images. This is not only important for the state estimation directly but also for the extraction algorithm supplying linear boundary features from pixelized images.

References

Aquilano, D., Pastero, L., Bruno, M., Rubbo, M., 2009. {1 0 0} and {1 1 1} forms of the NaCl crystals coexisting in growth from pure aqueous solution. Journal of Crystal Growth 311, 399-403.

Bajinca, N., de Oliveira, V., Borchert, C., Raisch, J., Sundmacher, K., 2010. Optimal control solutions for crystal shape manipulation. Computer Aided Chemical Engineering 28, 751-756.

Barnard, A.S., 2009. Shape-dependent confinement of the nanodiamond band gap. Crystal Growth & Design 9, 4860-4863.

Boerrigter, S.X.M., Josten, G.P.H., van de Streek, J., Hollander, F.F.A., Los, J., Cuppen, H.M., Bennema, P., Meekes, H., 2004. MONTY: Monte Carlo crystal growth on any crystal structure in any crystallographic orientation; application to fats. Journal of Physical Chemistry A 108, 5894-5902.

Borchert, C., Nere, N., Ramkrishna, D., Voigt, A., Sundmacher, K., 2009a. On the prediction of crystal shape distributions in a steady-state continuous crystallizer. Chemical Engineering Science 63, 686-696.

Borchert, C., Ramkrishna, D., Sundmacher, K., 2009b. Model based prediction of crystal shape distributions. Computer Aided Chemical Engineering 26, 141-146.

Briesen, H., 2006. Simulation of crystal size and shape by means of a reduced two-dimensional population balance model. Chemical Engineering Science 61, 104-112.

Calderon De Anda, J.A., Wang, X.Z., Lai, X., Roberts, K.J., 2005. Classifying organic crystals via in-process image analysis and the use of monitoring charts to follow polymorphic and morphological changes. Journal of Process Control 15, 785-797.

Castro, J.M., Cashman, K.V., Manga, M., 2003. A technique for measuring 3D crystal-size distributions of prismatic microlites in obsidian. American Mineralogist 88, 1230-1240.

Chakraborty, J., Singh, M.R., Ramkrishna, D., Borchert, C., Sundmacher, K., 2010. Modeling of crystal morphology distributions. Towards crystals with preferred asymmetry. Chemical Engineering Science 65 (21), 5676-5686.

Chemseddine, A., Moritz, T., 1999. Nanostructuring titania: control over nanocrystal structure, size, shape, and organization. European Journal of Inorganic Chemistry 2, 235-245.

Cole, A.J., 1966. Plane and stereographic projections of convex polyhedra from minimal information. Computer Journal 9, 27-31.

DeMenthon, D., Davis, L.S., 1992. Exact and approximate solutions of the perspective-three-point problem. IEEE Transactions on Pattern Analysis and Machine Intelligence 14 (11), 1100-1105.

Eggers, J., 2008. Modeling and Monitoring of Shape Evolution of Particles in Batch Crystallization Processes. Ph.D. Thesis. ETH Zurich.

Eilers, P.H.C., Goeman, J.J., 2004. Enhancing scatterplots with smoothed densities. Bioinformatics 20, 623-628.

Gilmer, G.H., 1980. Computer models of crystal growth. Science 208, 355-363.

Glicksman, M.E., Koss, M.B., Fradkov, V.E., Rettenmayr, M.E., Mani, S.S., 1994. Quantification of crystal morphology. Journal of Crystal Growth 137, 1-11.

Gonzalez, R.C., Woods, R.E., 2008. Digital Image Processing. Pearson Prentice Hall, New Jersey.

Hitec Zang, Germany. <http://www.hitec-zang.de/>.

Jerram, D.A., Mock, A., Davis, G.R., Field, M., Brown, R.J., 2009. 3D crystal size distributions: a case study on quantifying olivine populations in kimberlites. Lithos 112S, 223-235.

Kempkes, M., Eggers, J., Mazzotti, M., 2008. Measurement of particle size and shape by FBRM and in situ microscopy. Chemical Engineering Science 63, 4656-4675.

Kempkes, M., Vetter, T., Mazzotti, M., 2010a. Measurement of 3D particle size distributions by stereoscopic imaging. Chemical Engineering Science 65, 1362-1373.

Kempkes, M., Vetter, T., Mazzotti, M., 2010b. Monitoring the particle size and shape in the crystallization of paracetamol from water. Chemical Engineering Research & Design 88, 447-454.

Koster, A.J., Ziese, U., Verkleij, A.J., Janssen, A.H., de Jong, K.P., 2000. Three-dimensional transmission electron microscopy: a novel imaging and characterization technique with nanometer scale resolution for materials science. Journal of Physical Chemistry B 105, 7882-7886.

Larsen, P.A., Rawlings, J.B., 2009. The potential of current high-resolution imaging-based particle size distribution measurements for crystallization monitoring. AIChE Journal 55, 896-905.

Li, R.F., Thomson, G.B., White, G., Wang, X.Z., Calderon De Anda, J., Roberts, K.J., 2006. Integration of crystal morphology modeling and on-line shape measurement. AIChE Journal 52, 2297-2305.

Liu, X.Y., Boek, E.S., Briels, W.J., Bennema, P., 1995. Prediction of crystal growth morphology based on structural analysis of the solid-fluid interface. Nature 374, 342-345.

Lovette, M.A., Browning, A.R., Griffin, D.W., Sizemore, J.P., Snyder, R.C., Doherty, M.F., 2008. Crystal shape engineering. Industrial and Engineering Chemistry Research 47, 9812-9833.

Ma, D., Tafti, D., Braatz, R., 2002. High-resolution simulation of multidimensional crystal growth. Industrial and Engineering Chemistry Research 41, 6217-6223.

Ma, C.Y., Wang, X.Z., Roberts, K.J., 2008. Morphological population balance for modelling crystal growth in individual face directions. AIChE Journal 54, 209-222.

Mettler Toledo, USA. <www.mt.com.>.

Mullin, J.W., 2001. Crystallization. Butterworth-Heinemann, Oxford.

Myerson, A.S., 2002. Handbook of Industrial Crystallization. Elsevier, Oxford.

Patience, D.B., Rawlings, J.B., 2001. Particle-shape monitoring and control in crystallization processes. AIChE Journal 47, 2125-2130.

Persson, P.O., Strang, G., 2004. A simple mesh generator in MATLAB. SIAM Review 46 (2), 329-345.

Ramkrishna, D., 2002. Population Balances. Academic Press, San Diego.

Randolph, A.D., Larson, M.A., 1984. Theory of Particulate Processes. Academic Press, New York.

Retsch Technology, Germany. <http://www.retsch.de/>.

RJL Micro & Analytic GmbH, Germany. <http://www.rjl-microanalytic.com/>.

Selloni, A., 2008. Anatase shows its reactive side. Nature Materials 7, 613-615.

Sympatec GmbH, Germany. <www.sympatec.com>.

Taylor, J.E., Cahn, J.W., Handwerker, C.A., 1992. Geometric models of crystal growth. Acta Metallurgica et Materialia 40, 1443-1474.

Variankaval, N., Cote, A.S., Doherty, M.F., 2008. From form to function: crystallization of active pharmaceutical ingredients. AIChE Journal 54 (7), 1682-1688.

Wang, X.Z., Roberts, K.J., Ma, C., 2008. Crystal growth measurement using 2D and 3D imaging and the perspectives for shape control. Chemical Engineering Science 63, 1173-1184.

Weissbuch, L., Addadi, M., Leiserowitz, L., 1991. Molecular recognition at crystal interfaces. Science 253, 637-645.

Weissbuch, L., Popovitz-Biro, R., Lahav, M., Leiserowitz, L., 1995. Understanding and control of nucleation, growth, habit, dissolution and structure of two- and three-dimensional crystals using 'tailor-made' auxiliaries. Acta Crystallographica B 51, 115-148.

Weyland, M., Midgley, P.A., Thomas, J.M., 2001. Electron tomography of nanoparticle catalysts on porous supports: a new technique based on Rutherford scattering. Journal of Physical Chemistry B 105, 7882-7886.

Winn, D., Doherty, M.F., 2000. Modeling crystal shapes of organic materials grown from solution. AIChE Journal 46, 1348-1367.

Wintermantel, K., 1999. Process and product engineering achievements, present and future challenges. Chemical Engineering Science 54, 1601-1620.

Yang, H., Liu, Z.-H., 2009. Facile synthesis, shape evolution, photocatalytic activity of truncated cupros oxide octahedron microcrystals with hollows. Crystal Growth and Design 10, 2064-2067.

Yang, H.G., Sun, C.H., Qiao, S.Z., Zou, J., Liu, G., Campbell Smith, S., Cheng, H.M., Lu, G.Q., 2008. Anatase TiO2 single crystals with a large percentage of reactive facets. Nature 453, 638-641.

Yi, N., Si, R., Saltsburg, H., Flytzani-Stephanopoulos, M., 2009. Active gold species on cerium oxide nanoshapes for methanol steam reforming and the water gas shift reactions. Energy & Environmental Science 3, 831-837.