Direct Discretized Kernel Identification for Continuous Agglomeration Processes

Eric Otto [*] **Anton Maksakov** [*] **Robert Dürr** [***]
Stefan Palis [****] **Achim Kienle** [*,**]

[*] *Automation/Modelling, Otto von Guericke University Magdeburg, Universitätsplatz 2, 39106 Magdeburg, Germany*
[**] *Process Synthesis and Dynamics, Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstrasse 1, 39106 Magdeburg, Germany*
[***] *Engineering Mathematics, Magdeburg-Stendal University of Applied Sciences, Breitscheidstraße 2, 39114 Magdeburg, Germany*
[****] *National Research University Moscow Power Engineering Institute, Krasnokazarmennaya Ulitsa, 14, 111250 Moscow, Russia*

Abstract: Continuous particle agglomeration processes are important size-enlargement unit operations applied in the food, pharmaceutical and agricultural industry. For the improvement of these processes predictive mathematical models are of utmost importance. A widely applied modeling framework is the population balance equation, where the agglomeration kinetics are described by the so-called agglomeration kernel. The identification of functions describing these kinetics has turned out to be a challenging task. Therefore, this article deals with identifying such a kernel function by minimizing the L_2-residual between experimentally obtained particle size distributions and simulations. The application of a stochastic gradient descent algorithm with automatic differentiation for minimization allows for the direct identification of the high-dimensional matrix representing the discretized kernel function. The comparison between the simulated and measured size distribution shows that the identified kernel is able to accurately describe the evolution of the particle size distribution. The algorithm presented in this contribution can be applied to a variety of similar processes and the identified kernels can be used in process optimization and automation applications.

Keywords:, Process modeling and identification, Fluidized bed processes, Agglomeration processes, Agglomeration kernel, Optimization based identification

1. INTRODUCTION

Particle agglomeration is a size-enlargement process for solid particles, based on the aggregation of two or more particles, forming clusters. One typical technical realization is fluidized bed spray agglomeration (FBSA), which is a unit process widely used in the production of fertilizers, active pharmaceutical ingredients and various food powders (Bück and Tsotsas, 2016). During the FBSA process a particle bed is fluidized in an upwards faced air flow and the particle surface is wetted by a binder solution. After particle collision and drying of the binder new particles are formed (Fig. 1). The primary advantage of FBSA is the possibility of producing particles with predefined properties by adjusting certain process conditions. One of the most important product properties is the characteristic particle size, since it strongly affects other (mechanical) properties. Therefore the focus of this contribution lies on the particle volume v.

The particle production process can be improved in terms of efficiency and product quality by means of process intensification and automation, and especially process control (Otto et al., 2021a). For these purposes sound mathematical models are required, which therefore are an active field

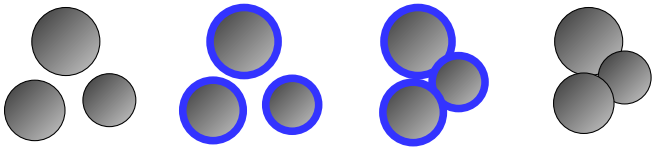

Fig. 1. Schematic representation of the 4-step wet agglomeration process.

of research. In the last decades the population balance framework has been established as a standard tool to describe the evolution of particulate systems (Ramkrishna, 2000). Therein, the process of binary agglomeration can be described by a partial differential equation where the kinetic rates are described by the so-called agglomeration kernel $\beta(t, u, u - v)$. Here, u and $(u - v)$ denote the volumes of particles forming an agglomerate with volume v and the value of beta describes the number of successful agglomeration events per unit of time and particle. Previous research has shown that the agglomeration ki-

netics depend on various process conditions (e.g. Strenzke et al. (2020)), therefore it is notoriously difficult to find meaningful mathematical descriptions. Some approaches based on the underlying micro-processes are presented in Ennis et al. (1991); Hussain et al. (2015), other researches have attempted to solve the inverse problem of finding kernel parameters from synthetic or actual measurement data (Chakraborty et al., 2015; Otto et al., 2021b) using generic function approaches. Due to the large amount of micro-processes involved in particle agglomeration which are themselves difficult to model, the former approach often results in complex model structures with unknown parameters. In contrast to this, the latter approach usually results in simple models with fixed structure and often a small number of tuning parameters. It is however an open challenge to find suitable approaches for the kernel functions.

This issue is circumvented by the non-parametric approach presented in this contribution. In the discrete version of the process PDE the kernel function reduces to a matrix, with the number of elements equal to the number of discretization classes squared. A variety of discrete formulations of the agglomeration PDE can be found in the literature, e.g. in Kumar et al. (2008); Singh et al. (2015). The identification of the matrix elements from experimental data can be formulated as an optimization problem, where the L_2-residual between the measured and a simulated particle size distribution is minimized as presented in Ramachandran and Barton (2010); Golovin et al. (2018); Otto et al. (2021b). In principle, the minimization can be conducted by applying any gradient descent method. In contrast to the contributions mentioned above the number of parameters is quite large if a matrix is identified, therefore computing the gradients by numerical differentiation becomes infeasible for high numbers of discretization classes. Alternatively, we use automatic differentiation in this contribution which is computationally more efficient and additionally allows for applications in online parameter estimation settings. Furthermore, this approach allows for other, high-parametric kernel function approaches such as neural networks. This has been presented in Nielsen et al. (2020) for general particle processes and the example of a flocculation/breakage process. In order to illustrate the effectiveness of the kernel modeling algorithm, it is applied to a set of synthetic and a set of experimental particle size distribution measurements. The accuracy of the identified kernel models is then validated by comparing simulations against the measured data.

The outline of this contribution is as follows: In section 2 the continuous process model as well as its discretization are described briefly. The kernel identification algorithm is presented in section 3, followed by results on synthetic and experimental data in section 4. A short conclusion and outlook are given in section 5.

2. PROCESS MODEL

An established framework for the modeling of agglomeration processes is the population balance equation (PBE), which describes the evolution of the number density distribution of particles $n(t, \mathbf{x})$ over time t and some in- or external coordinates $\mathbf{x}$. The focus in this contribution lies on the particle volume, i.e. $\mathbf{x} = v$, therefore the population balance is given by

$$\frac{\partial n(t,v)}{\partial t} = \dot{n}_\mathrm{f}(t,v) - \dot{n}_\mathrm{o}(t,v) + \dot{n}_\mathrm{a}(t,v) \tag{1}$$

as presented in Otto et al. (2021b). The left-hand side of Eq. (1) represents the accumulation of particles, the terms on the right-hand side account for the particle feed, outlet and agglomeration, which are described in more detail in the following. The particle feed

$$\dot{n}_\mathrm{f}(t,v) = \dot{N}_\mathrm{f}\, q_{0,\mathrm{feed}}(v) \tag{2}$$

is computed from the number feed rate $\dot{N}_\mathrm{f}$ and the normalized primary particle number density $q_{0,\mathrm{feed}}(v)$ which is assumed to be constant with respect to time and approximated by a normalized Gaussian function with mean value v_f and standard deviation σ_f. The particle outlet

$$\dot{n}_\mathrm{o}(t,v) = KT(v)n(t,v) \tag{3}$$

consists of the withdrawal rate K and the separation function $T(v)$ acting linearly on n. The separation function is approximated by a cumulative Gaussian function with mean value v_o and standard deviation σ_o

The agglomeration term is given by the following nonlinear integral (Hulburt and Katz, 1964; Ramkrishna, 2000):

$$\begin{aligned}\dot{n}_\mathrm{a}(t,v) = \frac{1}{2}\int_0^v \beta(t,u,v-u)n(t,u)n(t,v-u)\,\mathrm{d}u \\ - \int_0^\infty \beta(t,v,u)n(t,v)n(t,u)\,\mathrm{d}u\,,\end{aligned} \tag{4}$$

where $\beta(t,u,v)$ is the agglomeration kernel, describing the rate of successful agglomeration events between particles of volume u and v. Commonly the kernel function is separated into a time- and a size-dependent part

$$\beta(t,u,v) = \beta_0(t)\beta(u,v), \tag{5}$$

where $\beta_0(t)$ is the so-called agglomeration efficiency. The coalescence kernel, $\beta(u,v)$, is restricted to positive and symmetric functions due to simple physical considerations. In this contribution we furthermore assume β_0 to be constant. For a more detailed description of the process model we refer the reader to Otto et al. (2021b).

Since the desired kernel function is discretized and represented as a matrix, it is necessary to have a volume-discrete formulation of the population balance equation. Here, we apply the finite-volume discretization scheme developed in Singh et al. (2015), where the continuous volume range is partitioned into I_v classes and every class is represented by volume v_i. The total number of particles in class $i \in (1, ..., I_v)$ is denoted by $N_i(t)$ and can be approximated by

$$N_i(t) \approx n(t,v_i)\Delta v_i \tag{6}$$

with Δv_i being the width of class i. The discretized version of the agglomeration term (Eq. (4)) is given by

$$\frac{\mathrm{d}N_{\mathrm{a},i}}{\mathrm{d}t} = \frac{1}{2}\sum_{(j,k)\in\mathcal{I}^i} \beta_0\beta_{j,k}N_jN_kS_{i,j,k} - N_i\sum_j^{I_v}\beta_0\beta_{i,j}N_j \tag{7}$$

with

$$S_{i,j,k} = \frac{v_j + v_k}{v_i} \tag{8}$$

$$\mathcal{I}^i = \{(j,k) \in \mathbb{N}\times\mathbb{N} : v_{i-1/2} < (v_j + v_k) \leq v_{i+1/2}\}$$

for $i, j, k \in (1, ..., I_v)$. For further information regarding the discretization scheme, we refer to Singh et al. (2015).

In the discrete formulation the agglomeration kernel function $\beta(u, v)$ reduces to a matrix with I_v^2 elements $\beta_{i,j}$, which can be identified directly. Since the kernel function is symmetric, only the lower triangular elements are identified and then mirrored, resulting in identification $I_v(I_v - 1)/2 + 1$ variables. The discretization of the feed term and the outlet term is straightforward and not presented here.

3. KERNEL IDENTIFICATION

The kernel identification algorithm, presented schematically in Fig. 2, is based on the minimization of the L_2-residual between the measured and the simulated volume distribution which depends on the kernel function (Ramachandran and Barton, 2010; Golovin et al., 2018; Otto et al., 2021b). Starting at the given initial distribution, the population balance equation is solved numerically at the measurement time steps t_l and the following loss function is computed

$$J_{L_2}(p) = \frac{1}{2} \sum_{i=1}^{i=I_v} \sum_{l=1}^{l=I_t} \left(u^l_{i,\text{meas}} - u^l_{i,\text{sim}}(p) \right)^2 . \tag{9}$$

Here p denotes the vector of optimization variables, containing the kernel matrix elements $\beta_{j,k}$ and the agglomeration efficiency β_0. To compare experiments and simulations we use the diameter based volume distribution $u(t, d)$ for two reasons. Firstly, the experimental measurements are provided in this representation and secondly the diameter range of interest spans less orders of magnitude than the volume range, which is numerically advantageous.The volume distribution at diameter d_i and time $t = t_l$ is denoted by $u^l_i = u(t_l, d_i)$, which is computed by weighting the volume based number density distribution $n(t, v)$ with the volume and using the class-wise number conservation, i.e.:

$$u(t, d_i) = \frac{\pi}{6} d_i^3 n(t, v_i) \frac{\Delta v_i}{\Delta d_i}. \tag{10}$$

In the case with experimental measurement data $u(t, d)$ is recovered from the normalized volume density distribution $q_3(t, d)$ and the bed mass $m(t)$ (Otto et al., 2021b).

In a general agglomeration process, there will be volume classes without a significant amount of particles in it, hence the respective matrix elements $\beta_{j,k}$ of the kernel can not be identified. In order to distinguish these elements we introduce an additional, sparsity-promoting term to the cost functional

$$J_{L_1}(p) = \sum_{j,k} |\beta_{j,k}|, \tag{11}$$

rendering non-contributing matrix elements zero. The resulting loss function

$$J = \sum_{i,l} \left(u^l_{i,\text{meas}} - u^l_{i,\text{sim}}(p) \right)^2 + \gamma \sum_{j,k} |\beta_{j,k}| \tag{12}$$

with an additional tuning parameter γ is minimized using the stochastic gradient descent algorithm "Adam" (Kingma and Ba, 2017) implemented in python using the PyTorch library. The Adam algorithm uses parameter-specific adaption rates α in order to determine size of each parameter adaption step. Here, we choose different adaption rates for the matrix elements $\beta_{j,k}$ on the one hand and the agglomeration efficiency β_0 on the other. An advantage of using a stochastic method is the ability of minimizing objective functions with inherent stochasticity, in our case induced by the measurement noise. The discretized agglomeration term in the population balance equation tends to be stiff. Therefore using a Runge-Kutta 4(5) variable time step solver is advantageous. Direct backpropagation through the ODE-solver using automatic differentiation in backwards mode (Chen et al., 2018) allows to compute the necessary gradients efficiently, which is especially important regarding potential real-time applications.

The matrix elements $\beta_{j,k}$ are initialized with values, distributed randomly in the interval $[0, 1)$. The agglomeration rate, β_0 is initialized iteratively with suitable values. The non-negativity of the matrix elements is ensured by taking their absolute values before solving the population balance.

4. RESULTS

4.1 Matrix Element Identification from Synthetic Data

The parameter identification algorithm is applied to an exemplary set of synthetic measurement data, generated with the Kapur-Kernel presented in Fig. 3 (left) and additional process parameters presented in Tab. 1. The parameter estimation algorithm is terminated after 30 epochs due to the small value of the loss function and small gradient. The identified kernel matrix as well as the element-wise absolute error compared to the Kapur kernel are presented in the middle and on the right in Fig. 3. For volume classes with no or very few particle measured, here $v < 0.1$mm and $v > 1.5$mm, the corresponding matrix elements are close to zero due to the sparsity promoting L_1-term in the loss function. These regions are omitted in the graphical representation In regions with a significant number of particles measured, the difference between the kernels becomes small, but does not vanish.

The prediction capability of the identified kernel matrix is visualized and comparing the solution of the population balance equation to the measurements as presented in Fig. 4 for selected time instants. We see that the differences in

Table 1. Simulation and gradient descent parameters.

Parameter	**Value**	**Unit**
Feed		
$\dot{N}_\text{f}$	7.5×10^6	s^{-1}
μ_f	0.2	mm
σ_f	0.05	mm
Outlet		
K	3.6×10^{-3}	s^{-1}
μ_o	0.9	mm
σ_o	0.3	mm
Kernel		
β_0	3×10^{-10}	s^{-1}
a	0.5	
b	0.1	
Discretization		
I_v	40	
I_t	61	
Gradient descent		
α (β_0)	1×10^{-9}	
α ($\beta_{j,k}$)	1×10^{-2}	
γ	1×10^5	

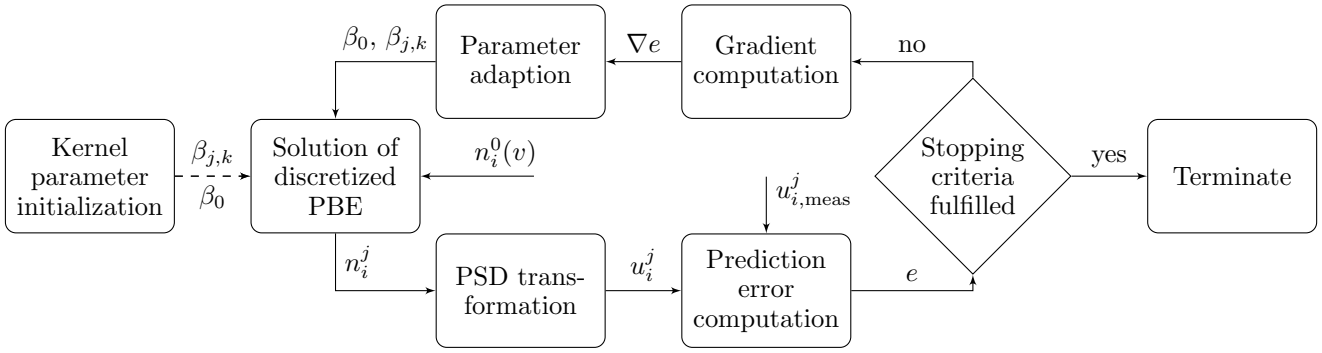

Fig. 2. Schematic presentation of the identification algorithm.

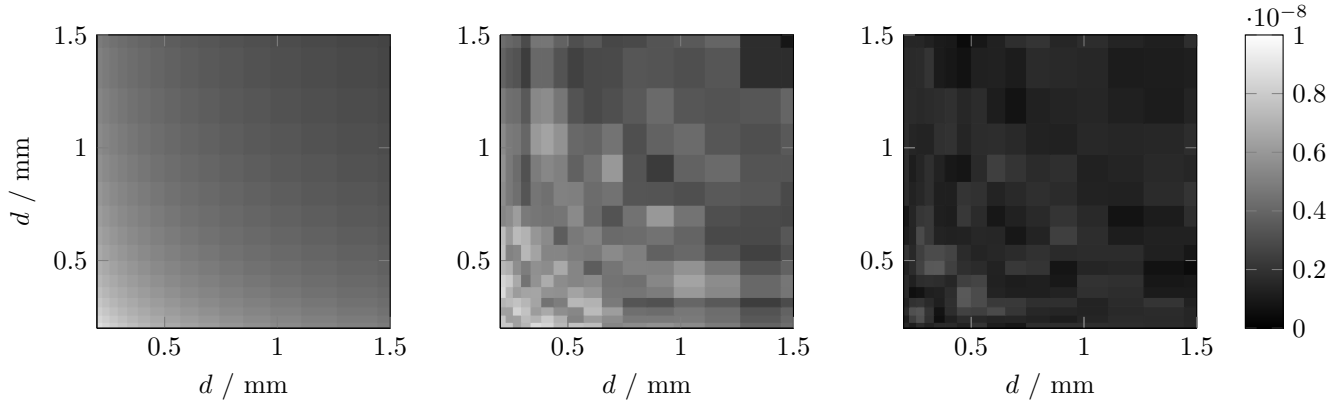

Fig. 3. Kapur kernel (left), identified matrix (middle) and their absolute difference (right).

the volume distribution nearly vanish, therefore the identified matrix clearly captures the relevant agglomeration kinetics. However, it should be clear that the kinetic rate constants of the nonlinear process can not be identified uniquely from one experiment alone and might even be non-unique for multiple experiments. Therefore, from a practical point of view, the kernel matrix should be identified from experiments covering the full range of initial conditions of interest.

4.2 Matrix Element Identification from Experimental Data

In this section the identification of a kernel matrix from actual experimental data is presented. The volume distribution measurements of a stable, continuous FBSA process were originally presented as "reference experiment" in Strenzke et al. (2020) and contain samples over the course of two hours under constant process conditions. The sample times of the original measurements were distributed non-uniformly. For the present paper the distributions were interpolated linearly at 61 linearly distributed time steps, resulting in measurement data every two minutes. In Otto et al. (2021b) Kapur kernel parameters were identified using the same data, however, only the dynamical behavior near the steady state could be described in good accuracy, i.e. the Kapur kernel model was not able to predict the measurements over the full two hour time horizon.

In order to compare the volume-based solution of the population balance equation with the diameter-based measured distributions they have to be converted into each other. Furthermore, we need values for K and T which change over the course of the experiment and can be recovered from the measurements of the volume distribution in the outlet and the bed mass over time. For the respective computations we refer the reader again to Otto et al. (2021b) and the references therein.

The measured volume distribution q_3 reaches a bi-modal steady state with a primary and a product particle peak after around one hour. Applying the parameter identification algorithm yields the kernel presented in Fig. 5 after 43 gradient descent steps which takes around 4 minutes on a Intel(R) Core(TM) I7-10610U CPU. The comparison between simulated and measured volume distribution in Fig. 6 shows high accuracy of the particle size prediction with a slight underestimation in the primary particle classes. Potential reasons for this could be overly simplifying model assumptions, such as lumping the spatial coordinates while in reality zone-formation occurs in the agglomeration chamber or the assumption of a time-independent kernel matrix.

In this contribution the evolution of the size distribution over the whole two hour time horizon was considered in the objective function, resulting in a kernel describing the dynamics over the same horizon. Note that, for some practical applications of a kernel identification algorithm, e.g. in model predictive control applications, a shorter prediction horizon with less computation time could be sufficient.

5. CONCLUSION AND OUTLOOK

In this contribution, the optimization-based identification of an agglomeration kernel for fluidized bed spray agglomeration processes was presented and validated. The application of automatic differentiation through an ODE-solver in a gradient descent method allowed the direct identification of 781 kernel matrix elements representing the kernel function. With the identified kernel matrices, the evolution of the particle size distributions could be predicted with high accuracy, both for synthetic and experimental measurement data sets.

Future research directions are manifold. Instead of identifying a kernel matrix for one experiment only, the identification algorithm can be applied to multiple data sets. By dividing these into training, test and validation data sets, the generalizability of the kernel matrix can be assessed. In this context, the kernel matrix can be replaced by any function approach (such as neural networks) in order to identify relationships between relevant process conditions such as temperature and gas moisture and the kernel. Since the algorithm is fast and can furthermore be parallelized, it can be applied parallel to the real process. Finally, the (online-) kernel identification algorithm can be embedded in any framework, depending on kernel identification, such as model based control algorithms.

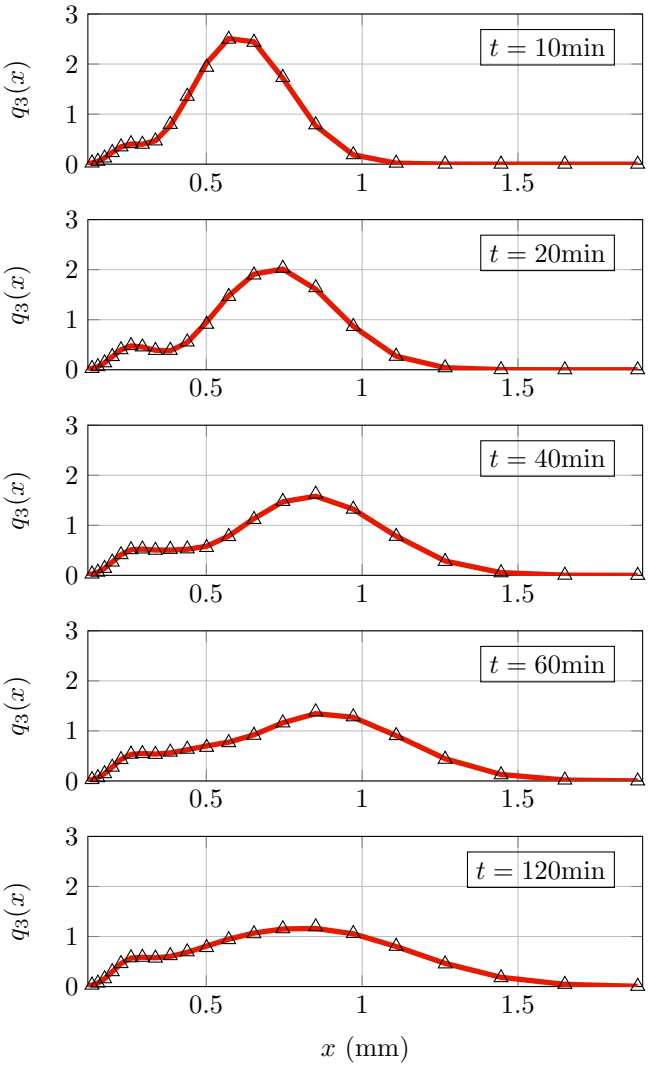

Fig. 4. Comparison between synthetic measurements and simulated particle size distributions.

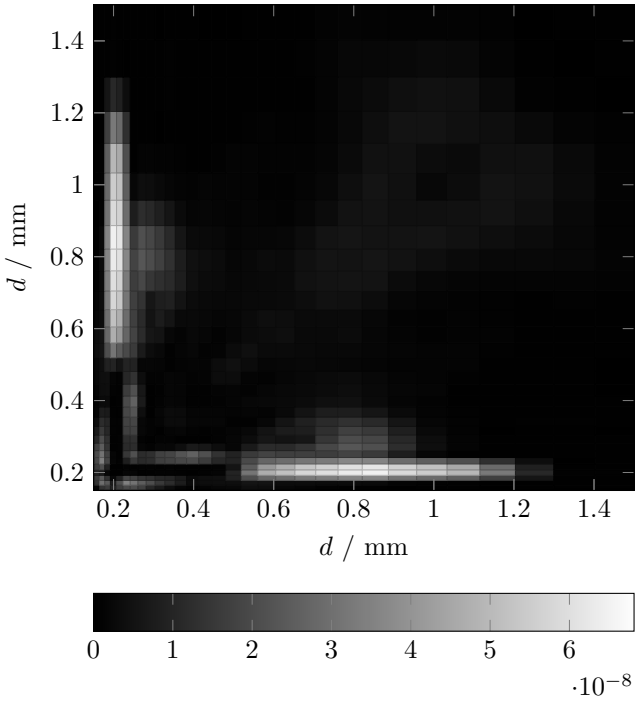

Fig. 5. Identified Matrix kernel for experimental data.

ACKNOWLEDGEMENTS

This work is funded by the European Regional Development Fund (ERDF) project "Center of Dynamic Systems". The financial support is hereby gratefully acknowledged.

REFERENCES

Bück, A. and Tsotsas, E. (2016). Agglomeration. In B. Caballero, P.M. Finglas, and F. Toldrá (eds.), *Encyclopedia of Food and Health*, 73 – 81. Academic Press, Oxford.

Chakraborty, J., Kumar, J., Singh, M., Mahoney, A., and Ramkrishna, D. (2015). Inverse problems in population balances. determination of aggregation kernel by weighted residuals. *Industrial & Engineering Chemistry Research*, 54(42), 10530–10538.

Chen, R.T.Q., Rubanova, Y., Bettencourt, J., and Duvenaud, D. (2018). Neural ordinary differential equations. *Advances in Neural Information Processing Systems*.

Ennis, B.J., Tardos, G., and Pfeffer, R. (1991). A microlevel-based characterization of granulation phenomena. *Powder Technology*, 65(1), 257 – 272.

Golovin, I., Strenzke, G., Dürr, R., Palis, S., Bück, A., Tsotsas, E., and Kienle, A. (2018). Parameter identification for continuous fluidized bed spray agglomeration. *Processes*, 6(12).

Hulburt, H. and Katz, S. (1964). Some problems in particle technology: A statistical mechanical formulation. *Chemical Engineering Science*, 19(8), 555 – 574.

Hussain, M., Kumar, J., and Tsotsas, E. (2015). A new framework for population balance modeling of spray fluidized bed agglomeration. *Particuology*, 19, 141–154.

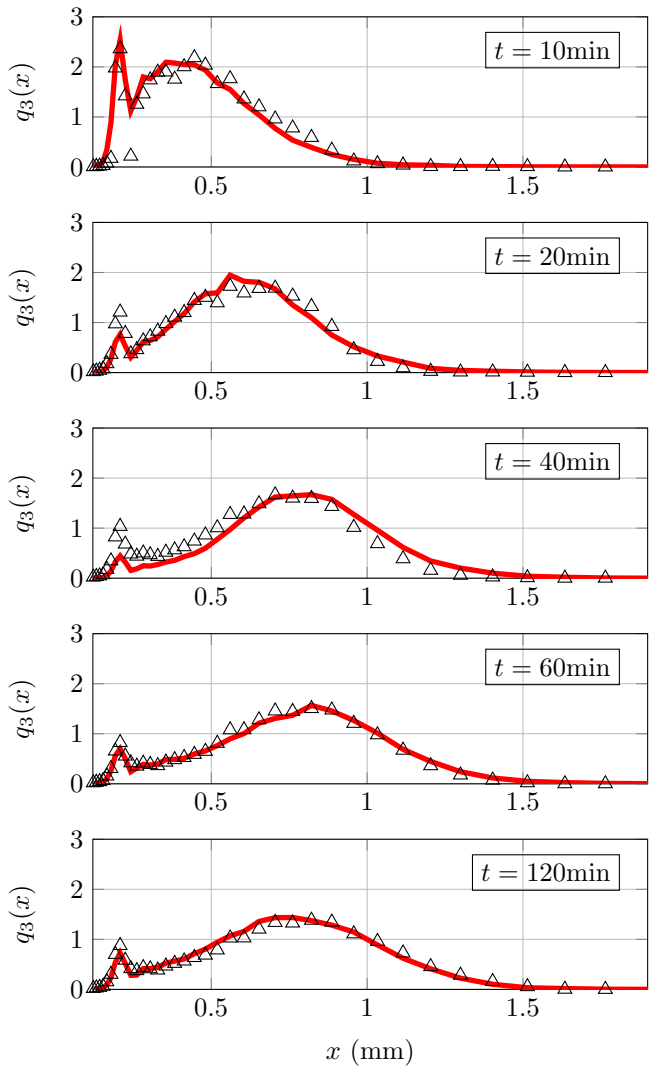

Fig. 6. Comparison between experimental measurements and simulated particle size distributions.

Kingma, D.P. and Ba, J. (2017). Adam: A method for stochastic optimization.

Kumar, J., Peglow, M., Warnecke, G., and Heinrich, S. (2008). The cell average technique for solving multi-dimensional aggregation population balance equations. *Computers & Chemical Engineering*, 32, 1810–1830.

Nielsen, R.F., Nazemzadeh, N., Sillesen, L.W., Andersson, M.P., Gernaey, K.V., and Mansouri, S.S. (2020). Hybrid machine learning assisted modelling framework for particle processes. *Computers & Chemical Engineering*, 140, 106916.

Otto, E., Behrens, J., Dürr, R., Palis, S., and Kienle, A. (2021a). Discrepancy-based control of particle processes. *Journal of Process Control*.

Otto, E., Dürr, R., Strenzke, G., Palis, S., Bück, A., Tsotsas, E., and Kienle, A. (2021b). Kernel identification in continuous fluidized bed spray agglomeration from steady state data. *Advanced Powder Technology*.

Ramachandran, R. and Barton, P.I. (2010). Effective parameter estimation within a multi-dimensional population balance model framework. *Chemical Engineering Science*, 65(16), 4884–4893.

Ramkrishna, D. (2000). *Population Balances: Theory and Applications to Particulate Systems in Engineering*. Academic Press.

Singh, M., Kumar, J., Bück, A., and Tsotsas, E. (2015). A volume consistent discrete formulation of aggregation population balance equation. *Mathematical Methods in the Applied Sciences*.

Strenzke, G., Dürr, R., Bück, A., and Tsotsas, E. (2020). Influence of operating parameters on process behavior and product quality in continuous spray fluidized bed agglomeration. *Powder Technology*, 375, 210 – 220.