


# Accounting for the Presence of Molecular Clusters in Modeling and Interpreting Nucleation and Growth

**Journal Article****Author(s):**

Ahn, Byeongho; [Bosetti, Luca](#) ; Mazzotti, Marco

**Publication date:**

2022-01-05

**Permanent link:**

<https://doi.org/10.3929/ethz-b-000519719>

**Rights / license:**

[Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International](#)

**Originally published in:**

Crystal Growth & Design 22(1), <https://doi.org/10.1021/acs.cgd.1c01193>

**Funding acknowledgement:**

788607 - Studying Secondary Nucleation for the Intensification of Continuous Crystallization (EC)

# Accounting for the Presence of Molecular Clusters in Modeling and Interpreting Nucleation and Growth

Published as part of a *Crystal Growth and Design virtual special issue in Celebration of the Career of Roger Davey*

Byeongho Ahn, Luca Bosetti, and Marco Mazzotti\*



Cite This: <https://doi.org/10.1021/acs.cgd.1c01193>



Read Online

ACCESS |



Metrics & More

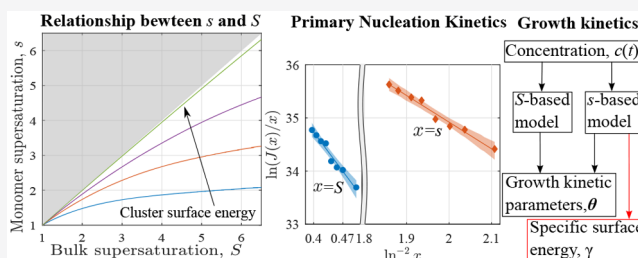


Article Recommendations



Supporting Information

**ABSTRACT:** The effect of molecular cluster formation on the estimation of kinetic parameters for primary nucleation and growth in different systems has been studied using computationally generated data and three sets of experimental data in the literature. It is shown that the formation of molecular clusters decreases the concentration of monomers and hence the thermodynamic driving force for crystallization, which consequently affects the crystallization kinetics. For a system exhibiting a strong tendency to form molecular clusters, accounting for cluster formation in a kinetic model is critical to interpret kinetic data accurately, for instance, to estimate the specific surface energy  $\gamma$  from a set of primary nucleation rates. On the contrary, for a system with negligible cluster formation, a consideration of cluster formation does not affect parameter estimation outcomes. Moreover, it is demonstrated that using a growth kinetic model that accounts for cluster formation allows the estimation of  $\gamma$  from typical growth kinetic data (i.e., desupersaturation profiles of seeded batch crystallization), which is a novel method of estimating  $\gamma$  developed in this work. The applicability of the novel method to different systems is proven by showing that the estimated values of  $\gamma$  are closely comparable to the actual values used for generating the kinetic data or the corresponding estimates reported in the literature.



## 1. INTRODUCTION

Nucleation and growth are fundamental crystallization phenomena in which solute molecules arrange into a crystal lattice. These two phenomena are typically driven by the volume and surface diffusion of solute molecules and their integration into the crystal lattice,<sup>1,2</sup> although in some nonclassical cases the crystal growth can be driven by the incorporation of dimers.<sup>3</sup>

Solute molecules in a solution comprise a mixture of monomers and molecular clusters, whose existence has been supported experimentally<sup>4–7</sup> by applying various methods and rationalized theoretically by using the classical nucleation theory (CNT)<sup>1</sup> and the two-step nucleation theory.<sup>8</sup> Accordingly, two types of solute concentrations can be used in defining the driving force for crystallization, namely, the bulk solute concentration,  $c$ , accounting for all solute molecules in the mixture, i.e., present both as monomers and in molecular clusters, and the monomer concentration,  $Z_1$ , considering only monomers. The bulk supersaturation  $S$  is defined as

$$S = \frac{c}{c_e} \quad (1)$$

where  $c_e$  is the bulk solubility (i.e., the bulk solute concentration at equilibrium). This supersaturation is often

used, partly because it is easier to characterize experimentally.<sup>9</sup> The monomer-based supersaturation  $s$  is defined as<sup>10</sup>

$$s = \frac{Z_1}{C_{1,e}} \quad (2)$$

where  $C_{1,e}$  is the monomer solubility (i.e., the monomer concentration at equilibrium).

These two types of supersaturations are equivalent under the assumption that all solute molecules in the liquid phase exist only as monomers, which overlooks the fact that part of the molecules can form molecular clusters. Consequently, to describe crystallization kinetics in the presence of the molecular clusters,  $s$  should be used instead of  $S$ . Note that crystallization kinetics described through these two supersaturations of reference can be closely comparable to each other; a similar discussion has also been presented elsewhere.<sup>9,11</sup>

**Received:** October 13, 2021

**Revised:** November 30, 2021

Despite the convenience of using the bulk supersaturation  $S$ , it is beneficial to use the monomer supersaturation  $s$  when inferring actual properties of a system, such as the specific surface energy,  $\gamma$ , from experimental data of primary nucleation rates, provided that the crystallization kinetics of the system is limited by the transport of monomers rather than that of oligomers. Given that  $\gamma$  is a key quantity in the study of several phenomena,<sup>12–19</sup> an accurate estimation of  $\gamma$  is essential. Moreover, by employing  $s$ , more information can be extracted from the same experimental data. For instance, by utilizing a growth kinetic model formulated on  $s$ ,  $\gamma$  can be estimated from growth kinetic data; this is a novel method of determining  $\gamma$  developed in this work. Note that a more profound understanding of primary nucleation and growth kinetics can offer a deeper insight into other crystallization phenomena, e.g., secondary nucleation, which is the dominant nucleation mechanism in industrial crystallization, particularly when operated continuously.<sup>20</sup>

This work has two objectives. The first is to demonstrate, for the first time, the effect of cluster formation on the estimation of crystallization kinetic parameters in different systems, with a focus on primary nucleation and growth. The second is to present a novel method of estimating  $\gamma$  exploiting the information embedded in growth kinetic data.

To achieve these objectives, it is crucial to acquire a set of comprehensive and unbiased data describing the kinetics of primary nucleation and growth. However, acquiring all data experimentally can be highly challenging. The nucleation rate can be easily under- or overestimated because of other concomitant mechanisms (e.g., aggregation) and/or technical challenges (e.g., an inadequate mixing of solutions and inherent limitations in available particle characterization devices).<sup>13</sup> A recent experimental investigation has also revealed that nucleation rates characterized by different methods can differ by several orders of magnitude.<sup>15</sup> Likewise, several issues can negatively affect the experimental characterization of growth kinetics. Seeded batch de-supersaturation experiments, which is a standard method of characterizing growth kinetics for process design,<sup>9,21–23</sup> can be accompanied by secondary nucleation.<sup>9,24</sup> In addition, some intrinsic properties of a compound (e.g., crystal morphology) can make measured data nonrepresentative. For instance, the particle size data of needle-like crystals measured by some 1D sizing techniques (e.g., laser diffraction or focused beam reflectance measurement) poorly represent the actual particle size distribution (PSD).<sup>25</sup>

To overcome these difficulties and to address the two objectives, measurement data are computationally generated using a first-principle mathematical model. This *in silico* approach prevents any potential systematic bias in data from confounding the present work; in silico approaches have been widely adopted in the literature, for instance, to study particle characterization techniques<sup>25–27</sup> and growth kinetic estimation.<sup>28</sup> In this contribution, data are generated using a kinetic rate equation (KRE) model. The generated data are first perturbed by Gaussian white noise to reflect measurement errors and then used to estimate the parameters of kinetic models for primary nucleation or for growth that are based on either  $s$  or  $S$ . Note that these kinetic models used for interpreting the data are completely independent of the KRE model used to generate the data. Estimated parameters and corresponding confidence intervals are compared with a set of the actual parameters characterizing a system described by the

KRE model, which reveals the conditions in which applying an  $s$ -based model (instead of a  $S$ -based one) is advantageous. Moreover, to prove the applicability of this analysis to real experimental data, three sets of experimental data in the literature are analyzed in the same manner.

This article is structured as follows. In section 2, the model and the methods for generating kinetic data are presented, and the three sets of reference experimental data are explained. In section 3, the methods for primary nucleation and growth kinetic modeling and parameter estimation are discussed. Finally, in section 4, the parameter estimation results and the novel method of estimating  $\gamma$  are presented and discussed.

## 2. ACQUISITION OF KINETIC MEASUREMENT DATA

The KRE model (see section 2.1) was used to computationally generate kinetic data for primary nucleation and growth individually (see section 2.2). Besides the generated data, three sets of experimental data in the literature (see section 2.3) were used to estimate the parameters of the kinetic models for primary nucleation and growth.

**2.1. Kinetic Rate Equation Model.** The conventional KRE model of nucleation (i.e., the Szilard model<sup>29</sup>) simulates primary nucleation, growth, and Ostwald ripening in a self-consistent manner in that all phenomena are explained by a unifying principle; all of them are described as a result of a series of molecule attachments to and detachments from molecular clusters.<sup>1,30</sup> This self-consistency is absent in the population balance equation (PBE) model where each phenomenon is explained by a distinctive kinetic/thermodynamic description. It is important to note that the description of the Szilard model is consistent with the CNT.<sup>1</sup> Although other nucleation mechanisms (e.g., the two-step nucleation theory) could be considered by adopting a corresponding kinetic description,<sup>31</sup> doing this is nevertheless beyond the scope of this work.

The KRE modeling framework has been successfully applied to describe several systems, e.g., the polymorphic system of silica,<sup>32</sup> that of L-glutamic acid,<sup>10</sup> and a system for the synthesis of semiconductor nanocrystals.<sup>33</sup> Moreover, the model has been extended to describe secondary nucleation caused by interparticle energies between seed crystals and molecular clusters,<sup>34</sup> with its description of underlying interparticle interactions built on a thermodynamic analysis.<sup>35</sup> In this work, the Szilard model was used for generating kinetic data. The model equations are<sup>1</sup>

$$\frac{dZ_n}{dt} = k_{n-1}^a Z_1 Z_{n-1} + k_{n+1}^d Z_{n+1} - k_n^a Z_1 Z_n - k_n^d Z_n \quad (n = 2, 3, \dots, n_{\max}) \quad (3a)$$

$$\frac{dZ_1}{dt} = - \sum_{n=2}^{n_{\max}} n \frac{dZ_n}{dt} \quad (3b)$$

where  $Z_n(t)$  is the concentration of molecular clusters consisting of  $n$  solute molecules (i.e., the  $n$ -sized clusters),  $t$  is the time,  $k_n^a$  and  $k_n^d$  are the rate constants of molecule attachment to and detachment from the  $n$ -sized clusters, respectively, and  $Z_{n_{\max}+1}(t) \equiv 0$ , with  $n_{\max}$  being the upper bound of the cluster size; this bound should be large enough to avoid any effect on numerical solutions.<sup>33</sup> The  $n$ -sized clusters represent crystals when the size  $n$  is larger than the critical nucleus size  $n^*$  defined in the CNT,<sup>1</sup> whereas they account for

part of the liquid phase when  $n$  is smaller than or equal to  $n^*$ , thus determining the bulk solute concentration  $c$ . Equation 3b is the mass balance for the solute molecules. The initial condition for eqs 3 is

$$Z_n(0) = Z_{n,0} \quad (4)$$

where  $Z_{n,0}$  is the initial concentration of the  $n$ -sized clusters. The expression for  $k_n^a$  is determined by the rate-limiting mechanism of the monomer transport to the  $n$ -sized clusters,<sup>1</sup> e.g., the volume-diffusion limited or surface-integration limited mechanism. In this work, the surface-integration limited mechanism is considered as an example since this mechanism is frequently encountered in practice.<sup>9,10,25,36</sup> A corresponding expression for  $k_n^a$  is<sup>1</sup>

$$k_n^a = \beta k_s d_1 D n^{2/3} \quad (5)$$

where  $\beta$  is the sticking coefficient,  $k_s$  is the surface shape factor,  $d_1 = (V_1/k_v)^{1/3}$  is the molecular diameter defined with the volume shape factor  $k_v$  and the molecular volume  $V_1$ , and  $D$  is the diffusion coefficient of the solute molecules in the solution. According to the assumption on which eq 5 is based, the clusters grow through the mechanism of rough growth.<sup>1,37–39</sup> The expression for  $k_n^d$  is<sup>1,34</sup>

$$k_n^d = k_{n-1}^a Z_1 \exp\left(\frac{\Delta G_n - \Delta G_{n-1}}{k_B T}\right) \quad (6)$$

where  $\Delta G_n$  ( $\Delta\mu$ ,  $\gamma$ ) is the Gibbs free energy for the formation of a  $n$ -sized cluster at a given driving force for crystallization,  $\Delta\mu$ ,  $k_B$  is the Boltzmann constant, and  $T$  is the absolute temperature of the system. With the initial condition (eq 4) and the rate constants (eqs 5 and 6), the system of the governing equations (eqs 3) can be solved to describe the temporal evolution of  $Z_{n,0}$ . To do that,  $\Delta G_n$  should be defined explicitly. Under the capillary approximation,<sup>1</sup> the Gibbs free energy  $\Delta G_n$  (in eq 6) is defined as

$$\Delta G_n = -n\Delta\mu + \gamma b n^{2/3} \quad (7)$$

where  $b = k_s(V_1/k_v)^{2/3}$  is the surface area of a solute molecule. The driving force  $\Delta\mu$  is the difference in the chemical potential between the solute molecule in the liquid phase and that in the solid phase, and it is given by<sup>10</sup>

$$\Delta\mu = k_B T \ln s \quad (8)$$

where  $s$  can be evaluated using eq 2 once the monomer solubility  $C_{1,e}$  (in eq 2) is known.

The monomer solubility  $C_{1,e}$  can be determined using the mass balance of the solute in the liquid phase, and the constraint that the bulk solute concentration  $c$  is equal to the bulk solubility  $c_e$  at phase equilibrium (i.e.,  $\Delta\mu = 0$ ). The mass balance in the liquid phase can be written as

$$c = \sum_{n=1}^{n^*} n Z_n \quad (9)$$

$$\approx \sum_{n=1}^{n^*} n C_n \quad (10)$$

where  $n^* = (2\gamma b/3\Delta\mu)^3$  is the critical nucleus size in the CNT<sup>1</sup> and the concentration  $Z_n$  is approximated by the equilibrium concentration of the  $n$ -sized clusters,  $C_n$ , considering that a population of subcritical and critical clusters reaches

equilibrium with a characteristic time shorter than those of nucleation and growth.<sup>30</sup> The equilibrium cluster concentration  $C_n$  is defined as<sup>1</sup>

$$C_n = C_0 \exp\left(-\frac{\Delta G_n}{k_B T}\right) \quad (11)$$

where  $C_0$  is the concentration of nucleation sites in the system on which the clusters of the solid phase can form.<sup>1</sup> The value of  $C_0$  can be determined by using eqs 7, 10, and 11, and the aforementioned constraint (i.e.,  $c = c_e$  when  $\Delta\mu = 0$ ):  $C_0 = c_e / \sum_{m=1}^{\infty} m \exp(-\Omega m^{2/3})$  where  $\Omega$  is the dimensionless surface energy defined as

$$\Omega = \gamma b / k_B T \quad (12)$$

One can determine  $C_{1,e}$  using eq 11 since  $C_{1,e}$  is by definition equal to  $C_1$  at  $\Delta\mu = 0$ .

The relationship between  $S$  and  $s$  can be derived in two steps: first, substituting eqs 7 and 8 into eq 11 yields

$$C_n = C_0 \exp(-\Omega n^{2/3}) s^n \quad (13)$$

Then, accounting for eqs 1, 12, and 13 in eq 10 leads to the following relationship between  $s$  and  $S$ :

$$S = \frac{\sum_{n=1}^{(2\Omega/3 \ln s)^3} n \exp(-\Omega n^{2/3}) s^n}{\sum_{m=1}^{\infty} m \exp(-\Omega m^{2/3})} \quad (14)$$

where  $\Omega$  is the only parameter.

The KRE model (eqs 2, 3–11) was solved by applying numerical methods widely used in the literature;<sup>30,32–34</sup> for small  $n$  (e.g.,  $n < 50$ ), a system of ordinary differential equations (ODEs) (eqs 3) can be integrated in time using an ODE solver (e.g., the MATLAB's ode15s), whereas, for large  $n$  (e.g.,  $n \geq 50$ ), the system of the ODEs can be adequately approximated by the Fokker–Planck partial differential equation, which can be solved efficiently using the weighted finite difference scheme proposed by Chang and Cooper.<sup>40</sup>

As shown in section S1 in the Supporting Information, nondimensionalization of the KRE model reveals that the general behavior of the model depends only on the initial condition and on the system-dependent property  $\Omega$ , thus suggesting that, for a given initial condition, the behavior of different systems can be studied by varying  $\Omega$ . To this aim, two values of  $\Omega = \{3, 10\}$  are used in this work. Concerning the other properties, those of the paracetamol–ethanol system at 20 °C (summarized in Table 1) are used as a reference.

**2.2. Data Generation Methods.** Primary nucleation kinetics are often determined based on the data of nucleation rates  $J$  characterized at different values of  $S$ .<sup>12,14–17,42</sup> To generate this type of data in silico, unseeded batch crystallization was simulated at different initial supersaturations using the KRE model, where the initial concentrations of subcritical clusters were set to zero; hence, only monomers were present at the start, which corresponds to a typical initial condition in the literature.<sup>30,33</sup> For each simulation, a nucleation rate,  $J$ , was calculated using the following definition:<sup>1</sup>

$$J = \left. \frac{d}{dt} \left( \sum_{n=n^*+1}^{n_{\max}} Z_n(t) \right) \right|_{t=t_s} \quad (15)$$

**Table 1. Physicochemical Properties of the Paracetamol–Ethanol System at 20 °C Used in the KRE Model**

properties	value	description
sticking coefficient, $\beta$	$3.2 \times 10^{-6}$	assumed <sup>a</sup>
surface-area shape factor, $k_s$	$\pi$	The spherical shape is assumed.
volume shape factor, $k_v$	$\pi/6$	The spherical shape is assumed.
molecular weight of solute, $M_w$	0.151 kg mol <sup>-1</sup>	
solid density, $\rho_s$	1263 kg m <sup>-3</sup>	
molar solubility, $c_M$	0.826 mol L <sup>-1</sup>	from ref 41
dynamic viscosity of solvent, $\eta$	$1.184 \times 10^{-3}$ Pa s	
molecular volume of solute, $V_1$	$1.99 \times 10^{-28}$ m <sup>3</sup>	$\equiv M_w/\rho_s N_A^b$
bulk solubility, $c_e$	$4.97 \times 10^{26}$ m <sup>-3</sup>	$\equiv 10^3 c_M N_A^b$
molecular diameter, $d_1$	$7.24 \times 10^{-10}$ m	$\equiv (V_1/k_v)^{1/3}$
diffusion coefficient, $D$	$5.01 \times 10^{-10}$ m <sup>2</sup> s <sup>-1</sup>	$\equiv k_B T/3\pi\eta d_1^c$

<sup>a</sup>This scales a nucleation rate and a growth rate such that their values are experimentally more relevant. <sup>b</sup> $N_A$  is the Avogadro number. <sup>c</sup>The Stokes–Einstein equation for diffusion in fluid.

Here,  $t_s$  is the time for the occurrence of stationary nucleation, when the monomer supersaturation  $s$ , the nucleation rate  $J$ , and the concentrations of subcritical clusters  $Z_{n<n^*}$  are practically time-independent;<sup>1</sup> after the phase of stationary nucleation, all the three quantities decline due to an appreciable consumption of monomers by a large population of growing clusters.<sup>1</sup> The bulk supersaturation  $S$  corresponding to  $J$  was calculated using eqs 1 and 10.

Growth kinetics is typically determined using the data of  $S$  profiles obtained from seeded batch de-supersaturation experiments with varying initial supersaturations.<sup>9,21–23</sup> To generate this type of data, seeded batch crystallization was simulated with different initial supersaturations using the KRE model, where the initial crystal size distribution follows a Gaussian distribution with a mean size of 50  $\mu\text{m}$  and a standard deviation of 5  $\mu\text{m}$ . The time evolution of  $S$  was calculated using eqs 1 and 10.

**2.3. Reference Experimental Data.** To show the general validity of this analysis, three sets of experimental data in the literature were analyzed in the same manner as the *in silico* data (see their main features in Table 2). The first set (E1)

**Table 2. List of Reference Experiments**

data	phenomenon	system	$T$ (K)	ref
E1	homogeneous nucleation	benzoic acid/sodium chloride/water	303.15	ref 14
E2	homogeneous nucleation	$\alpha$ -L-glutamic acid/water	298.15	ref 17
E3	growth	$\alpha$ -L-glutamic acid/water	298.15	ref 21

concerns homogeneous nucleation of benzoic acid from an aqueous sodium chloride solution at 30 °C,<sup>14</sup> the second set (E2) corresponds to homogeneous nucleation of  $\alpha$ -L-glutamic acid from an aqueous solution at 25 °C,<sup>17</sup> and the third set (E3) provides the growth kinetics of the same system at 25 °C.<sup>21</sup>

### 3. KINETIC PARAMETER ESTIMATION FOR PRIMARY NUCLEATION AND GROWTH

Kinetic models for primary nucleation and growth (see sections 3.1 and 3.2, respectively) were used to estimate model parameters in accordance with the procedure explained in section 3.3. This allowed for the interpretation of experimental data with and without consideration of cluster formation and its impact on crystallization kinetics. Note that the cluster formation can be considered in a kinetic model by using the monomer supersaturation  $s$  instead of the bulk supersaturation  $S$  as the supersaturation value of reference.

**3.1. Primary Nucleation Kinetics.** The primary nucleation rate  $J$  is typically described by using the following expression:<sup>1</sup>

$$J = \theta_i^{\text{PN}} x \exp\left(-\frac{\theta_2^{\text{PN}}}{\ln^2 x}\right) \quad (16)$$

where  $x \in \{s, S\}$  is a generic supersaturation and  $\theta_i^{\text{PN}}$  ( $i = 1, 2$ ) is the parameter to be estimated from a given set of primary nucleation rate data. Using the one-to-one correspondence between  $\theta_2^{\text{PN}}$  and  $\Omega$  according to the CNT (i.e.,  $\theta_2^{\text{PN}} = 4\Omega^3/27$ ),<sup>1</sup> corresponding values of  $\Omega$ ,  $s$  (eq 14), and  $\gamma$  (eq 12) can be determined from the estimated value of  $\theta_2^{\text{PN}}$ .

Equation 16 adequately approximates  $J$  obtained from the KRE model using eq 15, when  $x = s$ , and  $\theta_1^{\text{PN}}$  and  $\theta_2^{\text{PN}}$  are set as follows (see the relevant derivation in section S2 in the Supporting Information):

$$\theta_1^{\text{PN,KRE}} = \frac{\beta k_s d_1 D C_{1,e} C_0}{3} \sqrt{\frac{\Omega}{\pi}} \quad (17)$$

$$\theta_2^{\text{PN,KRE}} = \frac{4}{27} \Omega^3 \quad (18)$$

These parameters correspond to the actual nucleation kinetic parameters of a system described by the KRE model.

**3.2. Growth Kinetics. Population Balance Model.** The PBE model for a well-mixed batch reactor can be written as

$$\frac{\partial f}{\partial t} + G \frac{\partial f}{\partial L} = 0 \quad (19)$$

$$f(0, L) = f_0(L) \quad (20)$$

$$f(t, 0) = 0 \quad (21)$$

Here,  $L$  is the characteristic length of the crystal,  $G$  is the crystal growth rate assumed to be size-independent, and  $f(t, L)$  is the PSD, i.e.,  $f(t, L) dL$  being the number of crystals with length  $L \in [L; L + dL]$  per unit suspension volume.<sup>39</sup> Equation 20 is the initial condition, with  $f_0(L)$  being the initial PSD, while eq 21 is the boundary condition establishing that no nucleation occurs in the system. Note that neither agglomeration nor breakage is considered in the model. The PBE is coupled with the solute material balance

$$\frac{dc}{dt} = -\frac{k_v}{v_1} \frac{dm_3}{dt} \quad (22)$$

$$c(0) = c_0 \quad (23)$$

where  $c(t)$  is the solute concentration,  $m_i = \int_0^\infty L^i f(t, L) dL$  ( $i = 0, 1, 2, \dots$ ) is the  $i$ th moment of  $f$ ,  $c_0$  is the initial solute concentration, and  $v_1$  is the molecular volume.

The PBE was solved numerically by applying a fully discrete, high-resolution finite volume method with the van Leer Flux limiter,<sup>43</sup> while satisfying the convergence condition by Courant–Friedrichs–Lewy.<sup>44</sup> At the upper bound of the size domain, a numerical outflow boundary condition was imposed by using a zero-order extrapolation method.<sup>43</sup> For the discretization of  $f$  along  $L$ , a regular grid with a sufficiently small size was used to obtain a numerical solution independent of the grid size; in this work, the corresponding size ranged between 0.025 and 1  $\mu\text{m}$  depending on the data set used for fitting and a chosen growth model.

**Growth Rate Models.** Two growth rate models are considered in this work, namely, the rough growth (RG) model and the birth and spread (BS) model. The RG model was used for analyzing the data obtained from the KRE model whose description follows the rough growth mechanism, and a corresponding rate expression is<sup>1,37,39</sup>

$$G = \theta^{\text{RG}}(x - 1) \quad (24)$$

where  $\theta^{\text{RG}}$  is the parameter to be estimated. When  $x = s$  and a specific value of  $\theta^{\text{RG}}$  is assigned, the numerical solutions of the PBE model (eqs 19–24) approximate those of the KRE model that simulates seeded batch crystallization where nucleation is negligible, and the corresponding value of  $\theta^{\text{RG}}$  is as follows (see the relevant derivation in section S3 in the Supporting Information):

$$\theta^{\text{RG,KRE}} = \frac{\beta k_s d_1 D C_{1,e} \sqrt{v_1}}{3 \sqrt[3]{k_v}} \quad (25)$$

This equation gives the actual growth kinetic parameter of a system described by the KRE model.

The birth and spread growth model<sup>2</sup> was used when interpreting the experimental data E3 (see Table 2), considering that the growth kinetics of the corresponding system has been adequately described by the BS model in the literature.<sup>21,45</sup> The corresponding growth rate expression is<sup>2</sup>

$$G = \theta_1^{\text{BS}}(x - 1)^{2/3} (\ln x)^{1/6} \exp\left(-\frac{\theta_2^{\text{BS}}}{T^2 \ln x}\right) \quad (26)$$

where  $\theta_i^{\text{BS}}$  ( $i = 1, 2$ ) is the parameter to be estimated.

It is worth noting that using a  $s$ -based growth rate model allows the simultaneous estimation of the growth kinetic parameters and of the specific surface energy  $\gamma$ , which is a novel method of determining  $\gamma$ . The determination of  $\gamma$  is possible because, at a given solute concentration  $c$ , the growth rate  $G$  given by a  $s$ -based model depends not only on the kinetic parameters but also on  $\gamma$ , which affects the monomer supersaturation  $s$  (see eqs 1, 12, and 14). An application of this method to the generated data and to the experimental data E3 is presented in section 4.3.

**3.3. Fitting Procedure.** As detailed in section S4 in the Supporting Information, the parameter(s) of a kinetic model were estimated from a set of data points using the method of least-squares, and corresponding confidence intervals were estimated by applying a linear approximation of the model. Prior to parameter estimation, suitable model transformation and parameter transformation were applied, as explained in section S5 in the Supporting Information.

For nucleation kinetics, the logarithm of the nucleation rate ( $\ln J$ ) was chosen as a response variable in parameter estimation because it is the most common choice in the

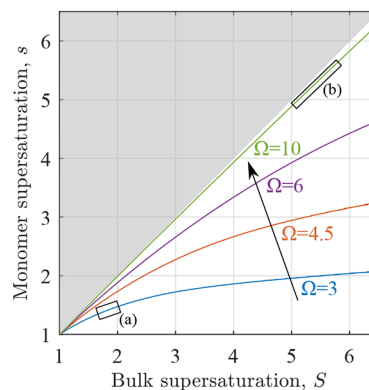
literature<sup>11,16,17,42</sup> and also because it linearizes the  $S$ -based model (eq 16 with  $x = S$ ) such that a simple linear regression can be applied. Moreover, the variance of  $\ln J$  in a set of experimental data is approximately constant,<sup>16,17</sup> suggesting that the logarithmic transformation is helpful to satisfy the constant-variance assumption in parameter estimation.<sup>46</sup> To reflect noise in experimental data, white Gaussian noise was added to the generated data of  $\ln J$  before using them for fitting.

For growth kinetics, the time evolution of bulk supersaturation  $S(t)$  was directly used as a response variable for parameter fitting, as it is common in the literature.<sup>21–23</sup> To give equal weight to each de-supersaturation profile regardless of its time duration, the profiles with a longer time duration were down-sampled such that the number of the data points of each profile used in parameter estimation is equal, which is a data preprocessing technique used also elsewhere.<sup>47</sup> To consider measurement noise in experimental data, white Gaussian noise was added to the generated  $S$  profiles before using them in parameter estimation, as done elsewhere.<sup>25,28</sup>

## 4. RESULTS AND DISCUSSION

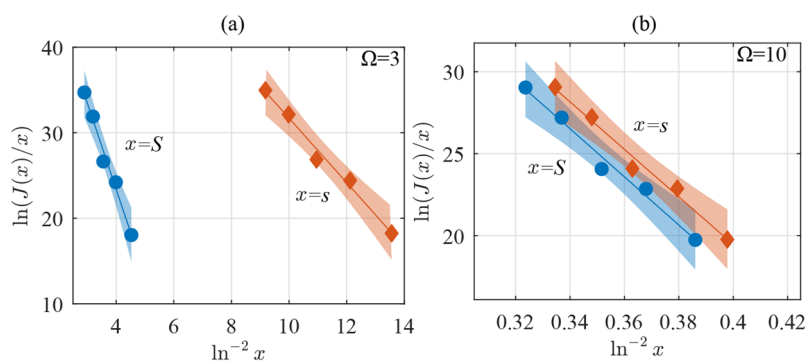
### 4.1. Comparison of Bulk and Monomer Supersaturation.

Figure 1 illustrates the relationship, given by eq 14,



**Figure 1.** Relationship between the bulk supersaturation  $S$  and the monomer supersaturation  $s$  at different values of the dimensionless surface energy  $\Omega = \{3, 4.5, 6, 10\}$ . The region (a) represents a part of the  $\Omega = 3$  system with a distinctive deviation between  $S$  and  $s$ , whereas the region (b) represents a part of the  $\Omega = 10$  system with a negligible difference between  $S$  and  $s$ . Gray area represents an unattainable region. Note that, for illustration purposes, the critical size  $n^*$  is treated as a continuously variable in this plot, thus yielding smooth curves for wide ranges of  $\Omega$ ,  $S$ , and  $s$ .

between the bulk supersaturation  $S$  and the monomer supersaturation  $s$  at different values of the dimensionless surface energy  $\Omega$ . The monomer supersaturation  $s$  is always smaller than or equal to the bulk supersaturation  $S$ , and their difference increases with an increase in  $S$  (or  $s$ ) and a decrease in  $\Omega$ . This is because both increasing  $s$  and decreasing  $\Omega$  lower the Gibbs free energy for cluster formation (see eqs 7, 8, and 12), which facilitates cluster formation while consuming the monomer concentration  $Z_1$  or, equivalently, the monomer supersaturation  $s$ . Accordingly, a system featuring a strong tendency to form molecular clusters exhibits a considerable deviation between  $s$  and  $S$  (e.g., see the region (a) in Figure 1). This deviation can cause different outcomes in kinetic parameter estimations for  $s$ -based and  $S$ -based models, whereas



**Figure 2.** Nucleation rate data generated with (a)  $\Omega = 3$  and (b)  $\Omega = 10$  used in parameter estimation (circle markers) for the  $s$ -based and  $S$ -based models, the corresponding ( $s, J$ ) data (diamond markers) where  $s$  is calculated using the fitted  $\theta_2^{\text{PN}}$  and eqs 14 and 18, the outputs of the fitted models (lines), and the corresponding 95% confidence intervals (shades).

a small deviation (e.g., the region (b) in Figure 1) would yield similar results, as demonstrated in sections 4.2 and 4.3 for primary nucleation kinetics and growth kinetics, respectively.

**4.2. Primary Nucleation Kinetics.** The effect of cluster formation on the estimation of primary nucleation kinetic parameters for the  $s$ -based and  $S$ -based models (i.e., eq 16 with  $x = s$  and  $x = S$ , respectively) is demonstrated in the following. Parameter estimation results based on a set of generated data corresponding to either region (a) or (b) in Figure 1, each representing a  $\Omega = 3$  or  $\Omega = 10$  system, respectively, are presented, followed by a discussion on the parameter estimation outcomes obtained from the experimental data E1 and E2.

**Generated Data.** Parameter estimation outcomes based on the  $\Omega = 3$  and  $\Omega = 10$  data are presented in Figure 2, panels a and b, respectively, where the nucleation rate data ( $S, J$ ), the corresponding ( $s, J$ ) data, the outputs of fitted models, and the corresponding 95% confidence intervals are reported.

In the case of  $\Omega = 3$ , the outcomes of the  $s$ -based and  $S$ -based models are distinctly different, thus clearly indicating that the interpretation of identical data can be strikingly different depending on the choice of the supersaturation of reference. Table 3 summarizes the actual value of primary nucleation parameters  $\theta^{\text{PN,KRE}}$  used to generate the  $\Omega = 3$  and  $\Omega = 10$  data, the fitted parameters for the  $s$ -based and the  $S$ -based model, denoted by  $\hat{\theta}^{\text{PN}}(x = s)$  and  $\hat{\theta}^{\text{PN}}(x = S)$ , respectively, and the corresponding 95% confidence intervals. For instance, the estimated value of  $\theta_1^{\text{PN}}$  for the  $s$ -based model

is  $1 \times 10^{30}$  with a 95% confidence interval ranging from  $8 \times 10^{26}$  to  $2 \times 10^{33}$ . It can be seen that, in the case of  $\Omega = 3$ , the actual parameter values can be satisfactorily estimated only when the  $s$ -based model is employed. On the contrary, using the  $S$ -based model underestimates  $\theta_1^{\text{PN}}$  and overestimates  $\theta_2^{\text{PN}}$ , which can be explained by considering the functional form of eq 16: since  $S$  is noticeably larger than  $s$  in the case of  $\Omega = 3$  (see Figure 1), a lower value of  $\theta_1^{\text{PN}}$  and a higher value of  $\theta_2^{\text{PN}}$  are necessary to obtain the same nucleation rate value. The overestimated  $\theta_2^{\text{PN}}$  yields an overestimation of  $\gamma$  (see eqs 12 and 18), which highlights that using the  $s$ -based model is necessary to estimate  $\gamma$  accurately for a system where cluster formation is significant.

Parameter estimation results obtained from the  $\Omega = 10$  data contrast sharply with those from the  $\Omega = 3$  data. In Figure 2b, the outputs of the two models fitted to the  $\Omega = 10$  data are closely comparable to each other, to the point that their 95% confidence intervals overlap. Moreover, as presented in Table 3, the fitted parameters for both models are extremely similar to the actual values, thus indicating that the two models are virtually equivalent in parameter estimation for a system where cluster formation is negligible.

**Experimental Data.** Parameter estimation outcomes for the data E1 and E2 are illustrated in Figure 3, panels a and b, respectively. In the case of E1 (corresponding to the benzoic acid/sodium chloride/water system as listed in Table 2), the behaviors of the two models are noticeably different, as previously observed for the  $\Omega = 3$  case. As summarized in Table 4, using the  $S$ -based model in parameter estimation yields a lower value of  $\theta_1^{\text{PN}}$  and a higher value of  $\theta_2^{\text{PN}}$  than those obtained using the  $s$ -based model, which is the same trend identified in the  $\Omega = 3$  case. According to the noticeable difference between two parameter estimation results based on the two models, it can be inferred that the system of E1 exhibits a strong tendency to form molecular clusters.

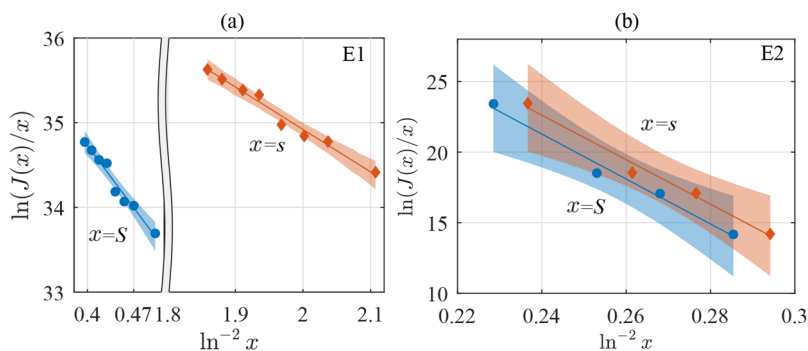
In Figure 3b, the results of parameter estimation using the data E2 (corresponding to the  $\alpha$ -L-glutamic acid/water system as listed in Table 2) shows characteristics observed in the  $\Omega = 10$  case. The 95% confidence intervals for the outputs of the two models overlap and the fitted parameter values for both models are extremely similar (see Table 4), thus suggesting that the formation of molecular clusters and its influence on parameter estimation can be regarded as insignificant in the system of E2.

**4.3. Growth Kinetics.** The aim of this section is to demonstrate the effect of cluster formation on the estimation

**Table 3.** Actual Parameters Used to Generate Two Data Sets ( $\theta^{\text{PN,KRE}}$ ) and Corresponding Parameter Estimation Results ( $\hat{\theta}^{\text{PN}}$ )<sup>a</sup>

data	parameters	$\theta_1^{\text{PN}}(\text{m}^{-3} \text{s}^{-1})$	$\theta_2^{\text{PN}}(-)$	$\gamma \times 10^3$ ( $\text{J m}^{-2}$ )
$\Omega = 3$	$\theta^{\text{PN,KRE}}$	$2 \times 10^{30}$	4.0	7.4
	$\hat{\theta}^{\text{PN}}(x = s)$	$1 \times 10^{30}$ [ $8 \times 10^{26}$ , $2 \times 10^{33}$ ]	3.8 [1.6, 5.9]	7.2 [5.5, 8.4]
	$\hat{\theta}^{\text{PN}}(x = S)$	$4 \times 10^{27}$ [ $2 \times 10^{24}$ , $1 \times 10^{31}$ ]	10.1 [7.9, 12.2]	10.0 [9.3, 10.7]
$\Omega = 10$	$\theta^{\text{PN,KRE}}$	$1 \times 10^{34}$	148	24.6
	$\hat{\theta}^{\text{PN}}(x = s)$	$4 \times 10^{33}$ [ $4 \times 10^{30}$ , $5 \times 10^{38}$ ]	145 [111, 179]	24.4 [22.3, 26.2]
	$\hat{\theta}^{\text{PN}}(x = S)$	$2 \times 10^{33}$ [ $1 \times 10^{28}$ , $2 \times 10^{38}$ ]	147 [113, 180]	24.5 [22.5, 26.2]

<sup>a</sup> $[\omega_{\min}, \omega_{\max}]$ : a 95% confidence interval of an estimate is from  $\omega_{\min}$  to  $\omega_{\max}$ .



**Figure 3.** Nucleation rate data (a) E1 and (b) E2 used in parameter estimation (circle markers) for the  $s$ -based and  $S$ -based models, the corresponding  $(s, J)$  data (diamond markers) where  $s$  is calculated using the fitted  $\theta_2^{\text{PN}}$  and eqs 14 and 18, the outputs of the fitted models (lines), and the corresponding 95% confidence intervals (shade). The data E1 are reprinted with permission from ref 14. Copyright 2001 American Institute of Chemical Engineers (AIChE). The data E2 are reprinted with permission from ref 17. Copyright 2006 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

**Table 4. Parameter Estimation Results Using Experimental Data E1 and E2 (See Table 2)<sup>a</sup>**

data	parameters	$\theta_1^{\text{PN}}$ ( $\text{m}^{-3} \text{s}^{-1}$ )	$\theta_2^{\text{PN}}$ (-)	$\gamma \times 10^3$ ( $\text{J m}^{-2}$ )
E1	$\hat{\theta}^{\text{PN}}(x=s)$	$3 \times 10^{19}$ , [ $1 \times 10^{19}$ , $1 \times 10^{20}$ ]	5.0 [1.6, 8.5]	9.7 [6.5, 11.5]
	$\hat{\theta}^{\text{PN}}(x=S)$	$8 \times 10^{16}$ , [ $4 \times 10^{16}$ , $2 \times 10^{17}$ ]	10.5 [8.7, 12.3]	12.4 [11.6, 13.0]
E2	$\hat{\theta}^{\text{PN}}(x=s)$	$2 \times 10^{26}$ , [ $7 \times 10^{19}$ , $5 \times 10^{32}$ ]	158 [98, 217]	29.7 [25.3, 33.0]
	$\hat{\theta}^{\text{PN}}(x=S)$	$7 \times 10^{25}$ , [ $2 \times 10^{19}$ , $3 \times 10^{32}$ ]	159 [100, 218]	29.7 [25.5, 33.0]

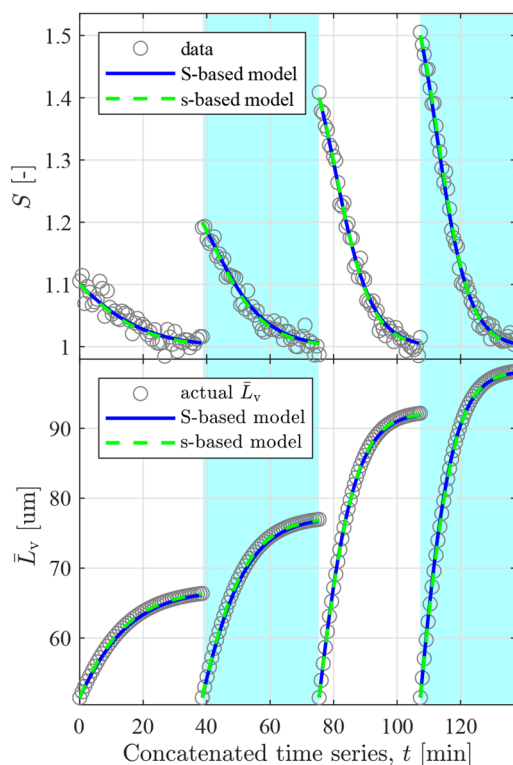
<sup>a</sup> $[\omega_{\min}, \omega_{\max}]$ : a 95% confidence interval of an estimate is from  $\omega_{\min}$  to  $\omega_{\max}$ .

of growth kinetic parameters and to present the novel method of estimating the specific surface energy  $\gamma$  from growth kinetic data. To this aim, parameter estimation outcomes based on two sets of data generated with  $\Omega = 3$  and  $\Omega = 10$  are analyzed, followed by a discussion on parameter estimation results using the experimental data E3.

**Generated Data.** In Figure 4, the concatenated time series of the  $\Omega = 3$  data and the corresponding predictions of the fitted,  $s$ -based, and  $S$ -based model (explained in section 3.2) are shown.

For both models, the only fitted output was the bulk supersaturation  $S$ , as explained in section 3.3, while the volume-weighted mean size  $\bar{L}_v$  was calculated according to its definition:  $\bar{L}_v = m_4/m_3$ . As shown in Figure 4, the outputs of both models agree with the corresponding data points satisfactorily, and, furthermore, the 95% confidence intervals for the outputs are narrower than the thickness of the lines corresponding to the model outputs, which means that the statistical goodness of fit of each model is adequate. It can be seen in Figure 4 that the time series of the mean size  $\bar{L}_v$  predicted by the two models overlap with the actual values obtained from the KRE simulations, thus suggesting that the descriptions of the two models are practically identical and these descriptions conform to that of the KRE model.

Although the outputs of the two models are remarkably similar, a striking difference is observed in the fitted values of the parameter ( $\theta^{\text{RG}}$ ), as summarized in Table 5. The value obtained using the  $S$ -based model is noticeably smaller than the actual one ( $\theta^{\text{RG,KRE}}$ ), whereas the one from the  $s$ -based model is closely comparable to the actual one. Recall that the



**Figure 4.** Concatenated time series of the data generated with  $\Omega = 3$  and the outputs of the fitted,  $s$ -based, and  $S$ -based models with respect to the bulk supersaturation  $S$  and the volume-weighted mean size  $\bar{L}_v$ . The alternating background color is used to distinguish a subset of the whole data, each generated from a different KRE simulation.

difference between the two models originates from the difference between  $s$  and  $S$  as illustrated in Figure 1. Moreover, by using the  $s$ -based model, the specific surface energy  $\gamma$  was simultaneously estimated as explained in section 3.2, which is the novel method of determining  $\gamma$ . Remarkably, the estimated value of  $\gamma$  ( $8.2 \text{ mJ m}^{-2}$ ) is very close to the actual one ( $7.4 \text{ mJ m}^{-2}$ ), which falls within the 95% confidence interval of the estimate ( $[6.0, 11.2] \text{ mJ m}^{-2}$ ), thus proving the applicability of the novel method to a system featuring a relatively small  $\Omega$ .

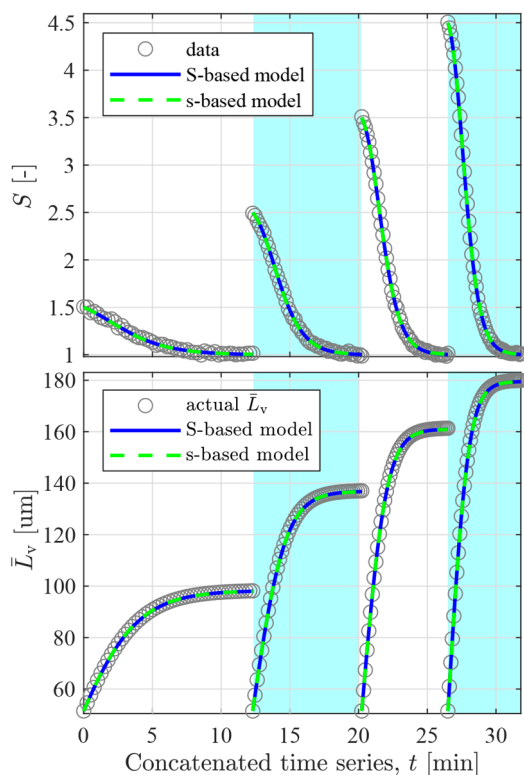
In Figure 5, the parameter estimation results based on the  $\Omega = 10$  data are shown. On the one hand, the results are similar to those for the  $\Omega = 3$  data in that the  $s$ -based and  $S$ -based



**Table 5. Actual Parameters in a Simulated Environment ( $\theta^{\text{RG,KRE}}$ ) and Parameter Estimation Results Using Two Sets of Generated Data ( $\hat{\theta}^{\text{RG}}$ )<sup>a</sup>**

data	parameters	$\theta^{\text{RG}} (\mu\text{m s}^{-1})$	$\gamma \times 10^3 (\text{J m}^{-2})$
$\Omega = 3$	$\theta^{\text{RG,KRE}}$	0.286	7.4
	$\hat{\theta}^{\text{RG}} (x = s)$	0.261 [0.189, 0.334]	8.2 [6.0, 11.2]
	$\hat{\theta}^{\text{RG}} (x = S)$	0.170 [0.167, 0.172]	
$\Omega = 10$	$\theta^{\text{RG,KRE}}$	0.435	24.6
	$\hat{\theta}^{\text{RG}} (x = s)$	0.436 [0.432, 0.440]	26.0 [23.2, 29.1]
	$\hat{\theta}^{\text{RG}} (x = S)$	0.430 [0.429, 0.431]	

<sup>a</sup>[ $\omega_{\min}$ ,  $\omega_{\max}$ ]: a 95% confidence interval of an estimate is from  $\omega_{\min}$  to  $\omega_{\max}$

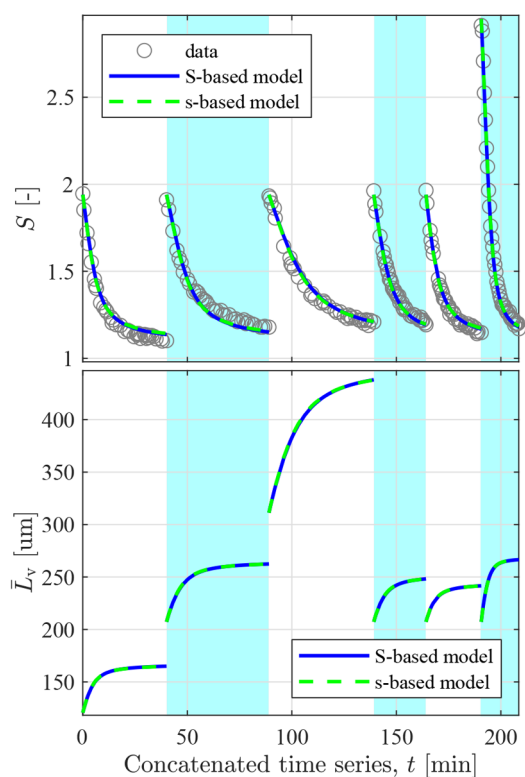


**Figure 5.** Concatenated time series of the data generated with  $\Omega = 10$  and the outputs of the fitted, *s*-based, and *S*-based models with respect to the bulk supersaturation *S* and the volume-weighted mean size  $\bar{L}_v$ . The alternating background color is used to distinguish a subset of the whole data, each generated from a different KRE simulation.

models fit the data satisfactorily: the outputs of the two models coincide with the data, and the corresponding 95% confidence intervals are narrower than the thickness of the lines representing the model outputs. On the other hand, the results differ from the ones obtained for the  $\Omega = 3$  data because, in the case of  $\Omega = 10$ , the estimated  $\theta^{\text{RG}}$  for the two models is practically the same, as summarized in Table 5. This is mainly because *s* and *S* are almost equivalent in the corresponding conditions, as shown in Figure 1. As reported in Table 5, the value of  $\gamma$  estimated from the  $\Omega = 10$  data (26.0  $\text{mJ m}^{-2}$ ) is closely comparable to the actual one (24.6  $\text{mJ m}^{-2}$ ), thus showing that the novel method of estimating  $\gamma$  is also applicable to a system exhibiting a relatively high  $\Omega$ .

**Experimental Data.** Finally, the parameter estimation results obtained using the experimental data E3 (corresponding to the  $\alpha$  L-glutamic acid/water system as listed in Table 2)

are illustrated in Figure 6. As explained in section 3.2, the birth and spread (BS) growth model (eq 26 with  $x = S$  or  $x = s$ ) was



**Figure 6.** Concatenated time series of the E3 data (see Table 2) and the outputs of the fitted, *s*-based, and *S*-based models with respect to the bulk supersaturation *S* and the volume-weighted mean size  $\bar{L}_v$ . The alternating background color is used to distinguish a subset of the whole data, each corresponding to a different experiment. From left to right, the fitted experiments correspond to Runs 1, 4, 5, 7, 10, and 12 in ref 21, which are all reported data sets measured at 25 °C, and they are reproduced with permission from the Royal Society of Chemistry.

utilized to fit the E3 data. In Figure 6, both *s*-based and *S*-based models fit data adequately, and the 95% confidence intervals for the predictions of the models are narrower than the width of the lines representing the predictions. Analogous to the results illustrated in Figures 4 and 5, the behavior of the two models illustrated in Figure 6 is essentially equivalent. Note that, unlike the BS model, the rough growth (RG) model could not describe the data E3 adequately (see section S6 in the Supporting Information). This clearly suggests that the growth of the  $\alpha$  L-glutamic acid crystals in aqueous solution is not governed by the RG mechanism, and thus the estimated  $\Omega$  value from the RG model cannot be trusted.

Table 6 summarizes the estimated parameters for the *s*-based and *S*-based models as well as those obtained from the reference literature (ref 21) for the E3 data, denoted by  $\hat{\theta}^{\text{BS}} (x = s)$ ,  $\hat{\theta}^{\text{BS}} (x = S)$ , and  $\hat{\theta}^{\text{BS,ref}}$ , respectively. There is a reasonable agreement between the estimated parameters for the *S*-based model  $\hat{\theta}^{\text{BS}} (x = S)$  and the reference values  $\hat{\theta}^{\text{BS,ref}}$  with a small difference in the value of  $\theta_2^{\text{BS}}$ . Note that it is not possible to obtain parameter values for the *S*-based model  $\hat{\theta}^{\text{BS}} (x = S)$  identical to the reference values  $\hat{\theta}^{\text{BS,ref}}$  because the applied parameter estimation methods are different, and the reference values  $\hat{\theta}^{\text{BS,ref}}$  were fitted using the whole data sets measured at 25, 35, and 45 °C, whereas only a subset measured at 25 °C

**Table 6. Parameter Estimation Results Using Experimental Data E3 (See Table 2)<sup>a</sup>**

data	parameters	$\theta_1^{\text{BS}}$ ( $\mu\text{m s}^{-1}$ )	$\theta_2^{\text{BS}} \times 10^{-4}$ ( $\text{K}^2$ )	$\gamma \times 10^3$ ( $\text{J m}^{-2}$ )
E3	$\hat{\theta}^{\text{BS,ref}}$	0.42	5.42	
	$\hat{\theta}^{\text{BS}}(x = s)$	0.57 [0.40, 0.74]	4.72 [4.44, 5.01]	14.2 [10.4, 19.4]
	$\hat{\theta}^{\text{BS}}(x = S)$	0.41 [0.40, 0.43]	4.97 [4.79, 5.14]	

<sup>a</sup> $[\omega_{\min}, \omega_{\max}]$ : a 95% confidence interval of an estimate is from  $\omega_{\min}$  to  $\omega_{\max}$ .

was employed for fitting the S-based model. The main motivation for using the 25 °C data only is to determine  $\gamma$ , a temperature-dependent quantity,<sup>10,16</sup> by applying the novel method and to compare its value with that estimated from the data E2, which is the primary nucleation kinetic data for the same system at the same temperature. It is worth mentioning that a comparison between the estimated parameters for the *s*-based and the *S*-based model shows that the estimated values of  $\theta_1^{\text{BS}}$  are noticeably different, and those of  $\theta_2^{\text{BS}}$  are highly comparable. The contrasting difference in the estimated  $\theta_1^{\text{BS}}$  suggests that the cluster formation in this system (i.e., an aqueous solution of L-glutamic acid) is likely to be significant.

The specific surface energy  $\gamma$  of  $\alpha$  L-glutamic acid crystal in an aqueous solution at 25 °C was estimated to be 29.7  $\text{mJ m}^{-2}$  in section 4.2 where the data E2 were used for fitting. However, this estimate of  $\gamma$  differs substantially from other four estimates reported in the literature,<sup>15,16</sup> namely, 6.8, 8.4, 10.4, and 17.1  $\text{mJ m}^{-2}$ , where two values of  $\gamma$  are given in ref 16, and the other two are calculated from the estimated values of  $\theta_2^{\text{N}}$  reported in ref 15 using eqs 12 and 18. A comparison of all the estimates based on five sets of nucleation rate data of the same system at the same temperature suggests that estimating  $\gamma$  from a set of nucleation rates is not highly reliable due to inherent difficulties in acquiring nucleation rates accurately (see ref 48 for a detailed discussion about inherent uncertainty in primary nucleation rate measurement due to its intrinsic stochasticity). On the contrary, the presented method of determining  $\gamma$  from the temporal evolution of a solute concentration (or bulk supersaturation) is intrinsically more robust, because the concentration can be monitored quite accurately using various techniques such as spectroscopy,<sup>49–52</sup> chromatography,<sup>53</sup> and particle imaging.<sup>52</sup> For instance, a concentration estimate obtained from attenuated total reflectance Fourier transform infrared (ATR-FTIR) spectroscopy can have an error of 3% with respect to the equilibrium concentration,<sup>54</sup> or even less.<sup>55</sup> As listed in Table 6, the value of  $\gamma$  estimated using our method is 14.2  $\text{mJ m}^{-2}$  (corresponding to  $\Omega = 4.9$ ), which is closely comparable to the two literature estimates (10.4 and 17.1  $\text{mJ m}^{-2}$ ) mentioned above. Although validating the presented method further by applying it to other systems is highly recommended, this is nevertheless outside the scope of this contribution.

The optimal method for estimating  $\gamma$  depends on a number of factors. If nucleation rates can be measured accurately, the standard method of fitting a nucleation rate model is more effective due to the relative simplicity of the fitting procedure. Moreover, the novel method can be inaccurate, when a quantitative estimation of solute concentration is difficult (e.g., due to a very low solubility<sup>54</sup>), when other concomitant mechanisms (e.g., breakage and secondary nucleation) occur uncontrollably during seeded batch de-supersaturation experiments, and when a system exhibits a very large value of  $\Omega$  such

that *s* and *S* are very similar. In the last case,  $\Omega$  cannot be accurately estimated because  $\Omega$  does not affect the growth rate through eq 14. When eq 14 reduces to  $S = s$ , then  $\Omega$  cannot be estimated from growth experiments.

## 5. CONCLUSIONS

The effect of molecular cluster formation on supersaturation and, consequently, on the estimation of kinetic parameters for primary nucleation and growth has been investigated in different systems using computationally generated data and three sets of benchmark experimental data. Utilizing the generated data has allowed us to study different systems under different conditions more extensively, to prevent any systematic error in experimental measurements from confounding our analysis, and to identify a general, system-dependent trend in cluster formation and its influence on the estimation of kinetic parameters. The generated data and the reference experimental data were used for fitting primary nucleation and growth kinetic models, while each model was formulated on the basis of supersaturation defined in two different manners. One supersaturation definition, denoted by *S*, was based on bulk concentrations, whereas another definition of supersaturation, denoted by *s*, was formulated based on monomer concentrations to consider the impact of cluster formation on kinetics. As a result, two kinetic models were formulated for each phenomenon.

Parameter estimation results using *s*-based and *S*-based models have demonstrated that considerable cluster formation in a system leads to strikingly different interpretations of identical data depending on the choice of the supersaturation of reference. In this case, an accurate estimation of model parameters can be achieved only when the *s*-based model is employed. On the contrary, when cluster formation is insignificant, its influence on parameter estimation outcomes is negligible.

A novel method of estimating the specific surface energy  $\gamma$  from concentration profiles of seeded batch crystallization has been developed. This method gives a more reliable estimate of  $\gamma$  than standard methods that utilize nucleation rate data, mainly because the concentration can be quantified more accurately than the nucleation rate. The applicability of the novel method to different systems has been proven by demonstrating that the estimates of  $\gamma$  agree with the actual values used to generate data or with the values reported in the literature.

It is worth noting that the present analysis is valid only within the framework of classical nucleation theory, where crystallization kinetics are limited by the transport of monomers instead of that of oligomers; molecular clusters have the same structure and specific surface energy as crystals and consist of molecules having the same bulk chemical potential as crystals independent of their cluster size. Although a generalization of this work is in principle possible, for instance, to be applicable to two-step nucleation theory, this is nonetheless beyond the scope of this contribution.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.cgd.1c01193>.

Nondimensionalization version of the Szilard model (section S1), derivation of the primary nucleation rate

expression (section S2), derivation of the growth rate expression (section S3), methods for estimating parameters and their confidence intervals (section S4), the transformation of the models and the parameters in parameter estimation (section S5), and the fitted outputs of the rough growth model to the experimental data E3 (section S6) (PDF)

## AUTHOR INFORMATION

### Corresponding Author

Marco Mazzotti – Institute of Energy and Process Engineering, ETH Zurich, 8092 Zurich, Switzerland; [orcid.org/0000-0002-4948-6705](https://orcid.org/0000-0002-4948-6705); Email: [marco.mazzotti@ipe.mavt.ethz.ch](mailto:marco.mazzotti@ipe.mavt.ethz.ch)

### Authors

Byeongho Ahn – Institute of Energy and Process Engineering, ETH Zurich, 8092 Zurich, Switzerland; [orcid.org/0000-0002-4405-5041](https://orcid.org/0000-0002-4405-5041)

Luca Bosetti – Institute of Energy and Process Engineering, ETH Zurich, 8092 Zurich, Switzerland; [orcid.org/0000-0001-6657-6510](https://orcid.org/0000-0001-6657-6510)

Complete contact information is available at: <https://pubs.acs.org/10.1021/acs.cgd.1c01193>

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

This project has received funding from the European Research Council (ERC) under the European Unions Horizon 2020 Research and Innovation Program under Grant Agreement No. 788607. The authors thank Mercedeh Sadat Hosseinalipour and Pietro Binel for critically reading the manuscript.

## DEDICATION

This is a contribution to the Virtual Special Issue of *Crystal Growth & Design* dedicated to Prof. Roger Davey and honoring his multifaceted pioneering contributions to crystallization science and technology. For the senior author of this work, Roger Davey has been, during the last two decades, a role model and a mentor, with his scientific insight and inspiring ideas, as well as with his kindness and friendship. His ideas have deeply influenced our understanding of the fundamentals of crystallization, particularly of nucleation and growth, including nucleation precursors in solution, and his books and papers have thoroughly shaped how we think of, speak about, and teach crystallization. We look forward to seeing him at crystallization conferences also after his retirement, accompanied by the smile that—the senior author is sure—neither Brexit nor COVID-19 have been able to turn off.

## NOMENCLATURE

TBD: units are different on a case-by-case basis

## ACRONYMS

BS birth and spread  
CNT classical nucleation theory  
KRE kinetic rate equation  
ODE ordinary differential equation  
PBE population balance equation  
PSD particle size distribution

RG rough growth

## GREEK SYMBOLS

$\beta$  sticking probability [–]  
 $\Delta\mu$  driving force for crystallization [J]  
 $\eta$  dynamic viscosity [Pa s]  
 $\gamma$  specific surface energy [J m<sup>-2</sup>]  
 $\Omega$  nondimensional surface energy [–]  
 $\omega$  arbitrary quantity [TBD]  
 $\rho_s$  solid density [kg m<sup>-3</sup>]  
 $\theta$  parameter [TBD]

## ROMAN SYMBOLS

$\bar{L}_v$  volume-weighted average size [ $\mu\text{m}$ ]  
 $b$  surface area of a solute molecule [ $\text{m}^2$ ]  
 $C$  equilibrium concentration of clusters at a given  $\Delta\mu$  [ $\text{m}^{-3}$ ]  
 $c$  bulk solute concentration [ $\text{m}^{-3}$ ]  
 $c_M$  molar solubility [ $\text{mol L}^{-1}$ ]  
 $C_0$  concentration of nucleation sites [ $\text{m}^{-3}$ ]  
 $D$  diffusion coefficient of a solute molecule in a solution [ $\text{m}^2 \text{s}^{-1}$ ]  
 $d_1$  molecular diameter [m]  
 $f$  particle size distribution [ $\text{m}^{-3} \mu\text{m}^{-1}$ ]  
 $G$  growth rate [ $\mu\text{m s}^{-1}$ ]  
 $J$  nucleation rate [ $\text{m}^{-3} \text{s}^{-1}$ ]  
 $k_B$  Boltzmann constant [J K<sup>-1</sup>]  
 $k_s$  surface area shape factor [–]  
 $k_v$  volume shape factor [–]  
 $k^a$  rate constant of molecule attachment [ $\text{m}^3 \text{s}^{-1}$ ]  
 $k^d$  rate constant of molecule detachment [ $\text{s}^{-1}$ ]  
 $L$  characteristic length of a crystal [ $\mu\text{m}$ ]  
 $M_w$  molecular weight [ $\text{kg mol}^{-1}$ ]  
 $m_i$   $i$ th moment of the particle size distribution  $f$  [ $\mu\text{m}^i \text{m}^{-3}$ ]  
 $n$  cluster size in terms of the number of solute molecules [–]  
 $N_A$  Avogadro number [ $\text{mol}^{-1}$ ]  
 $S$  supersaturation based on bulk solute concentrations [–]  
 $s$  supersaturation based on monomer concentrations [–]  
 $T$  absolute temperature [K]  
 $t$  time [s]  
 $t_s$  time for the occurrence of stationary nucleation [s]  
 $V$  volume [ $\text{m}^3$ ]  
 $v_1$  molecular volume in cubic micrometer [ $\mu\text{m}^3$ ]  
 $x$  generic supersaturation [–]  
 $Z$  concentration [ $\text{m}^{-3}$ ]  
 $\Delta G$  Gibbs free energy for forming a molecular cluster [J]

## SUPERSCRIPTS AND SUBSCRIPTS

0 initial value  
 $\wedge$  minimizer  
 $n$   $n$ -sized clusters  
\* critical nucleus  
BS birth and spread  
data data  
e phase equilibrium  
KRE kinetic rate equation  
max upper bound  
min lower bound  
PN primary nucleation  
ref reference  
RG rough growth

## REFERENCES

- (1) Kashchiev, D. *Nucleation*; Elsevier, 2000.
- (2) Ohara, M.; Reid, R. C. *Modeling Crystal Growth Rates from Solution*; Prentice-Hall: Englewood Cliffs, NJ, 1973.
- (3) Warzecha, M.; Verma, L.; Johnston, B. F.; Palmer, J. C.; Florence, A. J.; Vekilov, P. G. Olanzapine crystal symmetry originates in preformed centrosymmetric solute dimers. *Nat. Chem.* **2020**, *12*, 914–920.
- (4) Sleutel, M.; Lutsko, J.; van Driessche, A. E.; Durán-Olivencia, M. A.; Maes, D. Observing classical nucleation theory at work by monitoring phase transitions with molecular precision. *Nat. Commun.* **2014**, *5* (1), 5598.
- (5) Maes, D.; Vorontsova, M. A.; Potenza, M. A. C.; Sanvito, T.; Sleutel, M.; Giglio, M.; Vekilov, P. G. Do protein crystals nucleate within dense liquid clusters? *Acta Crystallogr., Sect. F: Struct. Biol. Commun.* **2015**, *71*, 815–822.
- (6) Yau, S.-T.; Vekilov, P. G. Quasi-planar nucleus structure in apoferritin crystallization. *Nature* **2000**, *406*, 494–497.
- (7) Chung, S.; Shin, S.-H.; Bertozzi, C. R.; De Yoreo, J. J. Self-catalyzed growth of S layers via an amorphous-to-crystalline transition limited by folding kinetics. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 16536–16541.
- (8) Vekilov, P. G. Dense liquid precursor for the nucleation of ordered solid phases from solution. *Cryst. Growth Des.* **2004**, *4*, 671–685.
- (9) Quilló, G. L.; Bhonsale, S.; Gielen, B.; van Impe, J. F.; Collas, A.; Xiouras, C. Crystal Growth Kinetics of an Industrial Active Pharmaceutical Ingredient: Implications of Different Representations of Supersaturation and Simultaneous Growth Mechanisms. *Cryst. Growth Des.* **2021**, *21*, 5403–5420.
- (10) Tahri, Y.; Kožisek, Z.; Gagnière, E.; Chabanon, E.; Bounahmidi, T.; Mangin, D. Modeling the Competition between Polymorphic Phases: Highlights on the Effect of Ostwald Ripening. *Cryst. Growth Des.* **2016**, *16*, 5689–5697.
- (11) Orlewski, P. M.; Mazzotti, M. Modeling of Mixing-Precipitation Processes: Agglomeration. *Chem. Eng. Technol.* **2020**, *43*, 1029–1039.
- (12) Blandin, A. F.; Mangin, D.; Nallet, V.; Klein, J. P.; Bossoutrot, J. M. Kinetics identification of salicylic acid precipitation through experiments in a batch stirred vessel and a T-mixer. *Chem. Eng. J.* **2001**, *81*, 91–100.
- (13) Roelands, C. P. M.; ter Horst, J. H.; Kramer, H. J. M.; Jansens, P. J. Analysis of Nucleation Rate Measurements in Precipitation Processes. *Cryst. Growth Des.* **2006**, *6*, 1380–1392.
- (14) Ståhl, M.; Åslund, B. L.; Rasmuson, Å. C. Reaction crystallization kinetics of benzoic acid. *AIChE J.* **2001**, *47*, 1544–1560.
- (15) Cedeno, R.; Maosoongnern, S.; Flood, A. Direct Measurements of Primary Nucleation Rates of p-Aminobenzoic Acid and Glutamic Acid and Comparison with Predictions from Induction Time Distributions. *Ind. Eng. Chem. Res.* **2018**, *57*, 17504–17515.
- (16) Lindenberg, C.; Mazzotti, M. Effect of temperature on the nucleation kinetics of  $\alpha$ -l-glutamic acid. *J. Cryst. Growth* **2009**, *311*, 1178–1184.
- (17) Schöll, J.; Vicum, L.; Müller, M.; Mazzotti, M. Precipitation of L-Glutamic Acid: Determination of Nucleation Kinetics. *Chem. Eng. Technol.* **2006**, *29*, 257–264.
- (18) Omar, W.; Mohnicke, M.; Ulrich, J. Determination of the solid liquid interfacial energy and thereby the critical nucleus size of paracetamol in different solvents. *Cryst. Res. Technol.* **2006**, *41*, 337–343.
- (19) Lindenberg, C.; Mazzotti, M. Continuous precipitation of L-asparagine monohydrate in a micromixer: Estimation of nucleation and growth kinetics. *AIChE J.* **2011**, *57*, 942–950.
- (20) Bosetti, L.; Mazzotti, M. Population Balance Modeling of Growth and Secondary Nucleation by Attrition and Ripening. *Cryst. Growth Des.* **2020**, *20*, 307–319.
- (21) Schöll, J.; Lindenberg, C.; Vicum, L.; Brozio, J.; Mazzotti, M. Precipitation of  $\alpha$ -l-glutamic acid: determination of growth kinetics. *Faraday Discuss.* **2007**, *136*, 247.
- (22) Mitchell, N. A.; Ó'Ciardhá, C. T.; Frawley, P. J. Estimation of the growth kinetics for the cooling crystallisation of paracetamol and ethanol solutions. *J. Cryst. Growth* **2011**, *328*, 39–49.
- (23) Garside, J.; Mersmann, A.; Nývlt, J. *Measurement of Crystal Growth and Nucleation Rates*; IChemE, 2002.
- (24) de Souza, B.; Cogoni, G.; Tyrrell, R.; Frawley, P. J. Evidence of Crystal Nuclei Breeding in Laboratory Scale Seeded Batch Isothermal Crystallization Experiments. *Cryst. Growth Des.* **2016**, *16*, 3443–3453.
- (25) de Albuquerque, I.; Mazzotti, M.; Ochsenein, D. R.; Morari, M. Effect of needle-like crystal shape on measured particle size distributions. *AIChE J.* **2016**, *62*, 2974–2985.
- (26) Borchert, C.; Sundmacher, K. Morphology evolution of crystal populations: Modeling and observation analysis. *Chem. Eng. Sci.* **2012**, *70*, 87–98.
- (27) Schorsch, S.; Ochsenein, D. R.; Vetter, T.; Morari, M.; Mazzotti, M. High accuracy online measurement of multidimensional particle size distributions during crystallization. *Chem. Eng. Sci.* **2014**, *105*, 155–168.
- (28) Salvatori, F.; Binell, P.; Mazzotti, M. Efficient assessment of combined crystallization, milling, and dissolution cycles for crystal size and shape manipulation. *Chem. Eng. Sci. X* **2019**, *1*, 100004.
- (29) Farkas, L. Keimbildungsgeschwindigkeit in übersättigten Dämpfen. *Z. Phys. Chem.* **1927**, *125U*, 236–242.
- (30) Vetter, T.; Iggländ, M.; Ochsenein, D. R.; Hänseler, F. S.; Mazzotti, M. Modeling Nucleation, Growth, and Ostwald Ripening in Crystallization Processes: A Comparison between Population Balance and Kinetic Rate Equation. *Cryst. Growth Des.* **2013**, *13*, 4890–4905.
- (31) Kashchiev, D.; Vekilov, P. G.; Kolomeisky, A. B. Kinetics of two-step nucleation of crystals. *J. Chem. Phys.* **2005**, *122*, 244706.
- (32) Ozkan, G.; Ortoleva, P. A mesoscopic model of nucleation and Ostwald ripening/stepping: Application to the silica polymorph system. *J. Chem. Phys.* **2000**, *112*, 10510–10525.
- (33) Rempel, J. Y.; Bawendi, M. G.; Jensen, K. F. Insights into the Kinetics of Semiconductor Nanocrystal Nucleation and Growth. *J. Am. Chem. Soc.* **2009**, *131*, 4479–4489.
- (34) Ahn, B.; Bosetti, L.; Mazzotti, M. Secondary Nucleation by Interparticle Energies. II. Kinetics *Cryst. Growth Des.* **2021**, DOI: 10.1021/acs.cgd.1c00928
- (35) Bosetti, L.; Ahn, B.; Mazzotti, M. Secondary Nucleation by Interparticle Energies. I. Thermodynamics *Cryst. Growth Des.* **2021**, DOI: 10.1021/acs.cgd.1c00927
- (36) Eder, C.; Choszcz, C.; Müller, V.; Briesen, H. Jamin-interferometer-setup for the determination of concentration and temperature dependent face-specific crystal growth rates from a single experiment. *J. Cryst. Growth* **2015**, *426*, 255–264.
- (37) Davey, R. J.; Garside, J. *From Molecules to Crystallizers: An Introduction to Crystallization*; Oxford University Press: New York, 2000.
- (38) Tilbury, C. J.; Doherty, M. F. Modeling layered crystal growth at increasing supersaturation by connecting growth regimes. *AIChE J.* **2017**, *63*, 1338–1352.
- (39) Randolph, A. D.; Larson, M. A. *Theory of Particulate Processes*; Elsevier, 1988.
- (40) Chang, J.; Cooper, G. A practical difference scheme for Fokker-Planck equations. *J. Comput. Phys.* **1970**, *6*, 1–16.
- (41) Jiménez, J. A.; Martínez, F. Temperature Dependence of the Solubility of Acetaminophen in Propylene Glycol + Ethanol Mixtures. *J. Solution Chem.* **2006**, *35*, 335–352.
- (42) Kashchiev, D.; van Rosmalen, G. M. Review: Nucleation in solutions revisited. *Cryst. Res. Technol.* **2003**, *38*, 555–574.
- (43) LeVeque, R. J. *Finite Vol. Methods for Hyperbolic Problems*; Cambridge University Press, 2002.
- (44) Courant, R.; Friedrichs, K.; Lewy, H. Über die partiellen Differenzgleichungen der mathematischen Physik. *Math. Ann.* **1928**, *100*, 32–74.
- (45) Kitamura, M.; Ishizu, T. Growth kinetics and morphological change of polymorphs of L-glutamic acid. *J. Cryst. Growth* **2000**, *209*, 138–145.

(46) Bates, D. M.; Watts, D. G. *Nonlinear Regression Analysis and Its Applications*; Wiley Series in Probability and Statistics; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 1988.

(47) Bötschi, S.; Rajagopalan, A. K.; Morari, M.; Mazzotti, M. Feedback Control for the Size and Shape Evolution of Needle-like Crystals in Suspension. IV. Modeling and Control of Dissolution. *Cryst. Growth Des.* **2019**, *19*, 4029–4043.

(48) Maggioni, G. M.; Bosetti, L.; Dos Santos, E.; Mazzotti, M. Statistical analysis of series of detection time measurements for the estimation of nucleation rates. *Cryst. Growth Des.* **2017**, *17*, 5488–5498.

(49) Cornel, J.; Lindenberg, C.; Mazzotti, M. Quantitative Application of in Situ ATR-FTIR and Raman Spectroscopy in Crystallization Processes. *Ind. Eng. Chem. Res.* **2008**, *47*, 4870–4882.

(50) Simone, E.; Zhang, W.; Nagy, Z. K. Application of Process Analytical Technology-Based Feedback Control Strategies To Improve Purity and Size Distribution in Biopharmaceutical Crystallization. *Cryst. Growth Des.* **2015**, *15*, 2908–2919.

(51) Hu, Y.; Liang, J. K.; Myerson, A. S.; Taylor, L. S. Crystallization Monitoring by Raman Spectroscopy: Simultaneous Measurement of Desupersaturation Profile and Polymorphic Form in Flufenamic Acid Systems. *Ind. Eng. Chem. Res.* **2005**, *44*, 1233–1240.

(52) Bötschi, S.; Rajagopalan, A. K.; Morari, M.; Mazzotti, M. An Alternative Approach to Estimate Solute Concentration: Exploiting the Information Embedded in the Solid Phase. *J. Phys. Chem. Lett.* **2018**, *9*, 4210–4214.

(53) Yang, Y.; Zhang, C.; Pal, K.; Koswara, A.; Quon, J.; McKeown, R.; Goss, C.; Nagy, Z. K. Application of Ultra-Performance Liquid Chromatography as an Online Process Analytical Technology Tool in Pharmaceutical Crystallization. *Cryst. Growth Des.* **2016**, *16*, 7074–7082.

(54) Rajagopalan, A. K.; Dual, A. Projection Imaging System to Characterize Crystallization Processes: Design and Applications, Ph.D. thesis, ETH Zurich, 2019.

(55) Frawley, P. J.; Mitchell, N. A.; Ó'Ciardhá, C. T.; Hutton, K. W. The effects of supersaturation, temperature, agitation and seed surface area on the secondary nucleation of paracetamol in ethanol solutions. *Chem. Eng. Sci.* **2012**, *75*, 183–197.