Supporting Information:
Secondary Nucleation by Interparticle Energies.

III. Nucleation Rate Model.

Byeongho Ahn,∗,†,‡ Luca Bosetti,∗,†,‡ and Marco Mazzotti∗,†

†Institute of Energy and Process Engineering, ETH Zurich, 8092 Zurich, Switzerland
‡These authors contributed equally.

E-mail: ahn@ipe.mavt.ethz.ch; bosetti@ipe.mavt.ethz.ch; marco.mazzotti@ipe.mavt.ethz.ch
S1 Experiment Conducted at 250 and 300 rpm

According to various experimental studies\textsuperscript{S1–S4} and empirical secondary nucleation rate models,\textsuperscript{S5} the secondary nucleation rate increases with the stirring rate. This implies that increasing the stirring rate facilitates secondary nucleation, thus accelerating the consumption of supersaturation, for instance, in seeded batch crystallization where secondary nucleation occurs. Contrary to this current understanding of secondary nucleation, the opposite trend can be found in a set of experimental data of the benchmark study\textsuperscript{S6} collected at three different agitation rates; as shown in Figure S1, the supersaturation profile of the experiment conducted at 200 rpm decays much faster than that at 250 rpm, and this clearly signifies that increasing the stirring rate from 200 rpm to 250 rpm decreases the secondary nucleation rate significantly, which is counter-intuitive. Moreover, the reproducibility of the experiments conducted at 250 rpm and 300 rpm is not shown in the benchmark study\textsuperscript{S6} while that at 200 rpm is. For these reasons, two experiments, performed at 250 rpm and 300 rpm, respectively, were not used in our work for fitting secondary nucleation models.

Figure S1: Effect of stirring rate on the desupersaturation curves of the experiments reported in the benchmark study.\textsuperscript{S6} Copyright Elsevier (2012)
Nucleation and growth are fundamental crystallization phenomena in which solute molecules are integrated into a crystal lattice. These two phenomena involve the volume and surface diffusion of solute molecules and their incorporation into the crystal lattice. These solute molecules constitute a mixture of monomers and molecular clusters, whose presence has been verified experimentally by employing several techniques, and explained theoretically by using classical nucleation theory (CNT) and the two-step nucleation theory. Consequently, to define the driving force for crystallization, two kinds of solute concentrations can be used: the bulk solute concentration, $c$, considering all solute molecules in the mixture, i.e., present both as monomers and in molecular clusters, and the monomer concentration, $Z_1$, accounting for only monomers.

The bulk supersaturation $S$ is defined as

$$S = \frac{c}{c_e}, \quad (S1)$$

while the monomer supersaturation $s$ is given by eq 8. The relation between the monomer solubility $C_{1,e}$ and the bulk solubility $c_e$ are given by

$$C_{1,e} = \frac{c_e \exp (-\Omega)}{\sum_{m=1}^{\infty} m \exp(-\Omega m^{2/3})}. \quad (S2)$$

The relationship between $S$ and $s$ can be expressed as

$$S = \frac{\sum_{n=1}^{(2\Omega/3\ln s)^3} n \exp (-\Omega n^{2/3}) s^n}{\sum_{m=1}^{\infty} m \exp(-\Omega m^{2/3})} \quad (S3)$$

where $\Omega$ is the dimensionless surface energy defined in Section 2.3.1.

In Figure S2, the relationship between $s$ and $S$ given by eq S3 is illustrated at different levels of the dimensionless surface energy $\Omega$. It can be seen that both increasing the bulk
Figure S2: Relation between the bulk supersaturation $S$ and the monomer supersaturation $s$ at different levels of the dimensionless surface energy $\Omega = \{3, 4.5, 6, 10\}$. Grey area marks an unattainable region in the $(S, s)$ plane. The lines are obtained using equations reported in ref S14.

supersaturation $S$ and decreasing the surface energy $\Omega$ reduce the corresponding monomer supersaturation $s$. This is because both an increase in $S$ and a decrease in $\Omega$ lower the thermodynamic energy barrier for cluster formation and thus allow monomers to form more molecular clusters,\textsuperscript{S14} which reduces the monomer concentration $Z_1$ and consequently the monomer supersaturation $s$. Accordingly, to account for the presence of molecular clusters in modeling nucleation and growth, the monomer supersaturation $s$ should be used as the supersaturation of reference (see ref S14 for a detailed discussion).

S3  Parameter Estimation

The values of model parameters were estimated by employing the method of least squares under the standard assumptions on measurement errors; the errors are independent of each other and distributed normally, with zero mean and a constant value of variance.\textsuperscript{S15} Under these assumptions, the least-square estimator is equivalent to the maximum-likelihood estimator.\textsuperscript{S16}

The fundamental principle of the least-squares method is to determine the $P$-component
parameter vector $\mathbf{\theta} = (\theta_1, \theta_2, \ldots, \theta_P)^T$ that minimizes the sum of squared residuals or the objective function $H(\mathbf{\theta})$ defined as:

$$H(\mathbf{\theta}) = \sum_{i=1}^{N} \left( y_i^\text{data} - y_i(\mathbf{\theta}) \right)^2$$  \hfill (S4)

where $N$ is the total number of experimental data points, $y_i^\text{data}$ is the value of a measured quantity of the $i$th data point, and $y_i(\mathbf{\theta}) = y(a_i^\text{data}; \mathbf{\theta})$ is the corresponding prediction of a model using a parameter vector $\mathbf{\theta}$ and a predictor variable $a_i^\text{data}$. To find the minimizer $\hat{\mathbf{\theta}}$ (i.e., the parameter vector that minimizes the objective function $H(\mathbf{\theta})$), a local optimizer (MATLAB’s lsqnonlin) was used from several starting points (by applying MATLAB’s MultiStart).

For a given $\hat{\mathbf{\theta}}$, a confidence interval (CI) of the estimated parameter $\hat{\theta}_i$ ($i = 1, 2, \ldots, P$) was determined by employing the linear approximation method, as explained in the following: an estimate for the variance of the errors $\sigma^2$ involving $(N - P)$ degrees of freedom can be calculated from:

$$\sigma^2 = \frac{H(\hat{\mathbf{\theta}})}{N - P}.$$  \hfill (S5)

The Jacobian matrix of the model evaluated at $\mathbf{\theta} = \hat{\mathbf{\theta}}$, $\mathbf{U}$, is given by

$$\mathbf{U} = \frac{\partial \mathbf{y}(\mathbf{\theta})}{\partial \mathbf{\theta}} \bigg|_{\mathbf{\theta} = \hat{\mathbf{\theta}}}$$  \hfill (S6)

where $\mathbf{y}(\mathbf{\theta}) = (y_1(\mathbf{\theta}), y_2(\mathbf{\theta}), \ldots, y_N(\mathbf{\theta}))^T$ is the vector of the model predictions for all data points. The positive semi-definite parameter covariance matrix $\mathbf{W}$ can be estimated from:

$$\mathbf{W} = \sigma^2 \mathbf{U}^T \mathbf{U}.$$  \hfill (S7)

Finally, a $(1 - \alpha)\%$ CI for the estimated parameter $\hat{\theta}_j$ ($j = 1, 2, \ldots, P$) can be calculated from:

$$\hat{\theta}_j \pm q_{1-\alpha/2} \sqrt{\frac{1}{N-P} W_{jj}}$$  \hfill (S8)
where \( t_{1-\alpha/2}^{N-P} \) is the \((1 - \alpha/2)\) quantile of the Student’s t-distribution based on \((N - P)\) degrees of freedom and \( W_{jj} \) is the \(j\)th diagonal element of \( W \).

## S4 Nucleation Rate from KRE Simulations

For each KRE simulation, a nucleation rate, \( J_{\text{KRE}} \), was computed according to its definition:

\[
J_{\text{KRE}} = \frac{d}{dt} \left. \left( \sum_{n=n_{\text{eff}}+1}^{n_{\text{max}}} Z_n(t) \right) \right|_{t=t_s}
\]  

(S9)

where \( n_{\text{max}} \) is the upper bound of the cluster size considered in this work, \( t_s \) is the time for the occurrence of the stationary nucleation during which the following quantities are nearly time-independent: monomer supersaturation \( s \), the nucleation rate \( J \), and the concentrations of subcritical clusters \( Z_{n_n < n_{\text{eff}}} \).

To compare \( J_{\text{KRE}} \) with \( J_{\text{SNIPE}} \) at various settings in Section 3.1, the following conditions were investigated: the initial supersaturation \( s_0 = \{1.64, 1.87, 2.10, 2.34\} \), the intensity of the stabilization effect \( E_{\text{st}} = \{1, 1.1, 1.2\} \), the range of the stabilization effect \( l_{\text{st}} = \{0, 1, 10, 100\} \), and the mean size of normally distributed seed crystals \( \{12.5 \mu m, 25 \mu m, 50 \mu m, 100 \mu m\} \), with the coefficient of variation of 1%. The other parameters of the KRE model remained unchanged from the values used in part II (see Table 2 in part II).

## S5 Estimation of Primary Nucleation Kinetics

Primary nucleation kinetic parameters can be estimated by fitting a primary nucleation rate model to a set of nucleation rates characterized at different levels of supersaturation. To this aim, eq 7 can be used as a model equation that accounts for the presence of molecular clusters in a solution, but its use in combination with a set of nucleation rates of the studied system (paracetamol in ethanol) at 20°C in the literature generated numerical issues. To overcome these difficulties and to determine the relevant nucleation kinetics, eq 7
with the monomer supersaturation $s$ replaced by the bulk supersaturation $S$ was employed, as done elsewhere.\textsuperscript{S13-S24} Note that this model equation can be linearized with respect to $1/\ln^2 S$, thus allowing for linear regression.

In linear regression, two different sets of nucleation rates were considered to reflect the inherent uncertainty in primary nucleation rate measurement and consequently the uncertainty in the estimated nucleation kinetics (for a detailed discussion on the intrinsic uncertainty in nucleation rate measurement due to the stochasticity of primary nucleation, see ref \textsuperscript{S25}). The first data set concerns a series of nucleation rates directly measured from the varying optical transmissivity of a nucleating solution,\textsuperscript{S24} while the second set consists of the nucleation rates predicted by the nucleation kinetic model that was fitted in ref \textsuperscript{S24} by applying the method of Nyvlt.\textsuperscript{S26} Note that the nucleation kinetic models fitted by employing the method of Kubota\textsuperscript{S27} (e.g., in refs \textsuperscript{S28} and \textsuperscript{S24}) were not used in this work, because, for the studied system at 20 °C, these models underestimate primary nucleation kinetics noticeably.\textsuperscript{S24}

Figure S3: Two sets of nucleation rate data used for fitting (markers), the outputs of the fitted models (lines), and the corresponding 95% confidence intervals (shade). Circle markers represent a set of nucleation rates predicted by a nucleation kinetic model that was fitted in ref \textsuperscript{S24} by applying the method of Nyvlt.\textsuperscript{S26} Diamond markers indicate a set of nucleation rates directly measured from the varying transmissivity of a nucleating solution,\textsuperscript{S24} which is reprinted with permission from ref \textsuperscript{S24}. Copyright © 2018 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.
In Figure S3, the two sets of nucleation rate data, the outputs of fitted models, and the corresponding 95% confidence intervals are illustrated. Although completely different approaches were adopted to acquire these two sets of nucleation rates, the values of the nucleation rate from the two sets are highly comparable with each other to the extent that the confidence intervals for the outputs of the fitted models overlap. All the estimated parameter values are summarized in Table S1.

Table S1: Parameter Estimation Results from Two Sets of Primary Nucleation Rates

| data                      | ref   | $\hat{k}_0 \times 10^{41}$ (m$^3$ s$^{-1}$) | $\hat{\Omega}$ (−) |
|---------------------------|-------|-------------------------------------------|---------------------|
| direct measurement        | [S24] | 1.95 [1.16, 3.28]                         | 0.76 [0.56, 0.89]   |
| from transmissivity       |       |                                           |                     |
| predictions of the model  | [S24] | 1.24 [0.85, 1.80]                         | 0.84 [0.73, 0.92]   |
| fitted by using the method|       |                                           |                     |
| of Nyvlt                  |       |                                           |                     |

$^a[\omega_{\min}, \omega_{\max}]$: a 95% confidence interval of an estimate is from $\omega_{\min}$ to $\omega_{\max}$. 
References

(S1) Bosetti, L.; Mazzotti, M. Study of Secondary Nucleation by Attrition of Potassium Alum Crystals Suspended in Different Solvents. *Cryst. Growth Des.* **2020**, *20*, 2570–2577.

(S2) Yousuf, M.; Frawley, P. J. Experimental Evaluation of Fluid Shear Stress Impact on Secondary Nucleation in a Solution Crystallization of Paracetamol. *Cryst. Growth Des.* **2018**, *18*, 6843–6852.

(S3) Yousuf, M.; Frawley, P. J. Secondary Nucleation from Nuclei Breeding and Its Quantitative Link with Fluid Shear Stress in Mixing: A Potential Approach for Precise Scale-up in Industrial Crystallization. *Org. Process Res. Dev.* **2019**, *23*, 926–934.

(S4) Wang, M.-L.; Huang, H.-T.; Estrin, J. Secondary nucleation of citric acid due to fluid forces in a Couette flow crystallizer. *AIChE J.* **1981**, *27*, 312–315.

(S5) Xu, S.; Hou, Z.; Chuai, X.; Wang, Y. Overview of Secondary Nucleation: From Fundamentals to Application. *Ind. Eng. Chem. Res.* **2020**, *59*, 18335–18356.

(S6) 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.

(S7) Ohara, M.; Reid, R. C. *Modeling Crystal Growth Rates from Solution*; Englewood Cliffs, N.J., Prentice-Hall, 1973.

(S8) Kashchiev, D. *Nucleation*; Elsevier, 2000.

(S9) 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*, 5598.
(S10) 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.

(S11) Yau, S.-T.; Vekilov, P. G. Quasi-planar nucleus structure in apoferritin crystallization. *Nature* **2000**, *406*, 494–497.

(S12) 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.* **2010**, *107*, 16536–16541.

(S13) Vekilov, P. G. Dense liquid precursor for the nucleation of ordered solid phases from solution. *Cryst. Growth Des.* **2004**, *4*, 671–685.

(S14) Ahn, B.; Bosetti, L.; Mazzotti, M. Accounting for the Presence of Molecular Clusters in Modeling and Interpreting Nucleation and Growth. *Cryst. Growth Des.* **2022**, *22*, 661–672.

(S15) 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.

(S16) Seber, G. A. F.; Wild, C. J. *Nonlinear Regression*; Wiley Series in Probability and Statistics; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 1989.

(S17) Ahn, B.; Bosetti, L.; Mazzotti, M. Secondary Nucleation by Interparticle Energies. II. Kinetics. *Cryst. Growth Des.* **2022**, *22*, 74–86.

(S18) Kashchiev, D.; van Rosmalen, G. M. Review: Nucleation in solutions revisited. *Cryst. Res. Technol.* **2003**, *38*, 555–574.

(S19) Schöll, J.; Vicum, L.; Müller, M.; Mazzotti, M. Precipitation of L-Glutamic Acid: Determination of Nucleation Kinetics. *Chem. Eng. Technol.* **2006**, *29*, 257–264.

S-10
Lindenberg, C.; Mazzotti, M. Effect of temperature on the nucleation kinetics of α-l-glutamic acid. *J. Cryst. Growth* **2009**, *311*, 1178–1184.

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.

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.

Ståhl, M.; Åslund, B. L.; Rasmuson, Å. C. Reaction crystallization kinetics of benzoic acid. *AIChE J.* **2001**, *47*, 1544–1560.

Maosoongnern, S.; Flood, A. E. Validation of Models Predicting Nucleation Rates from Induction Times and Metastable Zone Widths. *Chem. Eng. Technol.* **2018**, *41*, 2066–2076.

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.

Nývlt, J. Kinetics of nucleation in solutions. *J. Cryst. Growth* **1968**, *3-4*, 377–383.

Kubota, N. A new interpretation of metastable zone widths measured for unseeded solutions. *J. Cryst. Growth* **2008**, *310*, 629–634.

Mitchell, N. A.; Frawley, P. J.; Ó'Ciardiá, C. T. Nucleation kinetics of paracetamol–ethanol solutions from induction time experiments using Lasentec FBRM®. *J. Cryst. Growth* **2011**, *321*, 91–99.