A. Solution of the Master equation

Define the probability generating function:[1]

$$\varphi(z,t) = \sum_{n=0}^{\infty} P_n(t)z^n = P_0(t) + P_1(t)z + P_2(t)z^2 + \cdots + P_n(t)z^n + \cdots \tag{S1}$$

Multiplying the Master equation (1) for each n by z^n and summing for all n results in a single differential equation

$$\frac{\partial \varphi(z,t)}{\partial t} = (z-1)\kappa(t)\varphi(z,t), \quad \varphi(z,0) = 1. \tag{S2}$$

Redefining the time scale by $\tau = \int_0^t \kappa(s)ds$ allows this equation to be written as

$$\frac{\partial \varphi(z,\tau)}{\partial \tau} = (z-1)\varphi(z,\tau), \quad \varphi(z,0) = 1, \tag{S3}$$

whose analytical solution is

$$\varphi(z,\tau) = e^{(z-1)\tau} = e^{(z-1)\int_0^t \kappa(s)ds} = e^{z\int_0^t \kappa(s)ds} e^{-\int_0^t \kappa(s)ds}. \tag{S4}$$

Expanding $e^{z\int_0^t \kappa(s)ds}$ in a Taylor series and matching coefficients gives the probabilities:

$$P_n(t) = \frac{1}{n!}\left[\int_0^t \kappa(s)ds\right]^n e^{-\int_0^t \kappa(s)ds}, \quad n = 0,1,2,\ldots, \tag{S5}$$

which was reported for $n = 0$ for crystallization in droplets by Koop et al.[2]

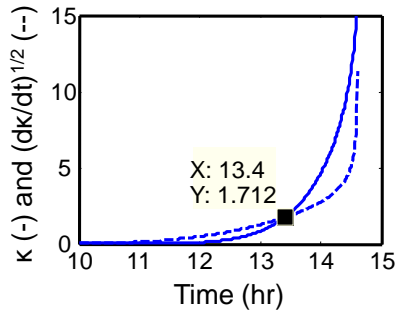## B. Example calculation of the maximum-likelihood induction time

Figure S1. Plot of $\kappa(t)$ and $\sqrt{d\kappa/dt}$ vs. time for Experiment Condition #1 in Table 1; the intercept is at the maximum likelihood value in Eq. (17).

C. Quantification of uncertainty in kinetic parameters

Confidence intervals on each model parameter are included here for completeness, although they provide a poor characterization of uncertainties when the parameters are highly correlated, such as occurs when estimating kinetics in crystallization processes.[3,4] Uncertainties in best-fit kinetic parameters $\beta^* = [A^* \ B^*]^{\mathrm{T}}$ were quantified using t-statistics and F-statistics.[5] For t-statistics the induction times were written in terms of a series expansion:

$$\tilde{t}_{ind}^i(\beta) = t_{ind}^i(\beta^*) + F_i(\beta - \beta^*) + \text{higher order terms}, \tag{S6}$$

where

$$F_i = \left. \frac{\partial \tilde{t}_{ind,i}}{\partial \beta} \right|_{\beta = \beta^*} \tag{S7}$$

is the Jacobian that was numerically computed using central differences. The perturbation used for computing the Jacobian was determined by plotting the central difference vs. magnitude of the perturbation. If the perturbation is too small, the Jacobian oscillates as perturbation varies; if the perturbation is too large, the Jacobian either increases or decrease as perturbation changes. The optimal perturbation is in between these two effects where the Jacobian remains more or less constant as perturbation changes (see Fig. S2). The parameter covariance matrix, V_β, is given by

$$V_\beta^{-1} = \sum_{i=1}^{n_{exp}} F_i^T V_y^{-1} F_i \tag{S8}$$

where the measurement error covariance matrix V_y is a diagonal matrix with

$$\sigma_y^2 = \frac{1}{\upsilon} \sum_{i=1}^{n_d} (t_{ind,i} - t_{ind,measured,i})^2 \tag{S9}$$

as diagonal entries, where $\upsilon = n_d - N_\beta$ is the number of degrees of freedom and n_d is the number of data points. The $100(1-\alpha)\%$ confidence interval for each parameter was determined from

$$\beta_i^* - t_{\alpha/2}(\upsilon)\sqrt{V_{\beta,ii}} \leq \beta_i \leq \beta_i^* + t_{\alpha/2}(\upsilon)\sqrt{V_{\beta,ii}} \tag{S10}$$

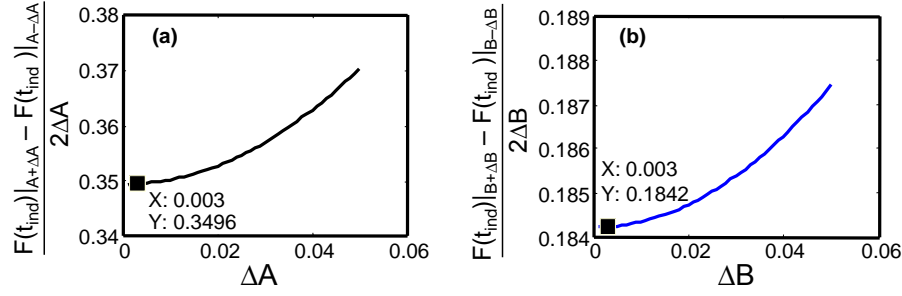where $t_{\alpha/2}$ is the t-distribution.

Figure S2. Jacobian computed by central differences vs. perturbation magnitude for (a) parameter *A* and (b) parameter *B* for Experimental Condition #2 in Table 2.

In *F*-statistics, gridding was performed to determine the parameters that give the sum-of-squared errors ψ within the bound ψ^*:

$$\psi \leq \left(1 + \frac{N_\beta F_\alpha (N_\beta, \upsilon)}{\upsilon}\right)\psi^* \tag{S11}$$

where F_α is the *F*-distribution. *F*-statistics takes into account nonlinearity but it is more computationally expensive than *t*-statistics, as several iterations are required to determine the range of parameters such that ψ satisfies the *F*-test.[5]

D. Details on Kolmogorov-Smirnov statistics for Section III.B

The test statistics for Experimental Conditions #1 and #2 are $d_{30} = 0.19$ and $d_{15} = 0.28$. For $\alpha = 0.1$ and $N = 30$, the critical value of Kolmogorov-Smirnov distribution is 0.22, hence the null hypothesis H_0 that $F(t)$ is the distribution (11) is not rejected for the data for Experimental Condition #1. For $\alpha = 0.1$ and $N = 15$, the critical value of Kolmogorov-Smirnov distribution is 0.30, hence H_0 is also not rejected for data set 2. The hypothesized distribution function is accepted as an appropriate model to describe the empirical distribution function at a significance level of $\alpha = 0.1$.

References

[1] Kendall, D. G. *J. Roy. Stat. Soc. Ser. B* **1949**, *11*, 230.
[2] Koop, T.; Luo, B. P.; Biermann, U. M.; Crutzen, P. J.; Peter, T. *J. Phys. Chem. A* **1997**, *101*, 1117.
[3] Miller, S. M.; Rawlings, J. B. *AIChE J.* **1994**, *40*, 1312.
[4] Ma, D. L.; Chung, S. H.; Braatz, R. D. *AIChE J.* **1999**, *45*, 1469.
[5] Beck, J. V.; Arnold, K. J. *Parameter Estimation in Engineering and Science*, Wiley: New York, 1977.