International Journal of Modern Physics B, Vol. 16, Nos. 1 & 2 (2002) 367–374

IDENTIFICATION OF KINETIC PARAMETERS IN MULTIDIMENSIONAL CRYSTALLIZATION PROCESSES

RUDIYANTO GUNAWAN, DAVID L. MA,
MITSUKO FUJIWARA, and RICHARD D. BRAATZ

Chemical Engineering, University of Illinois, 600 South Mathews Avenue, Box C-3 Urbana, Illinois 61801, United States of America

Received (8 June 2001)
Revised (15 September 2001)

Advances in sensor technology and increased competition in the pharmaceutical industry have generated significant interest in the identification of models for the solution formation of crystals with multiple characteristic dimensions. A procedure is proposed that uses a small number of batch experiments to identify the kinetic parameters for multidimensional crystallization processes. The parameters are estimated simultaneously from the on-line measurement of infrared spectra and from cross-moments of the crystal size distribution. The identification procedure maximizes the informativeness of the data produced by each experiment, produces an estimate of the accuracy of the kinetic parameters, and allows the consideration of competing hypotheses for characterizing the crystallization kinetics. The parameter identification strategy is applied to the batch crystallization of potassium dihydrogen phosphate, which forms two-dimensional crystals from solution. To the best of the authors' knowledge, this is the first time that the kinetic parameters for a multidimensional crystallization process are identified from a small number of batch experiments.

1. Introduction

Advances in sensor technology and increased competition in the pharmaceutical industry have generated significant interest in the characterization of multidimensional crystallization processes.[1] The crystallization processes considered here have large numbers of crystals produced from solution. New crystals may be formed directly from solution, by breakage of existing crystals, by attrition due to collisions, or by removal of a semi-ordered surface layer through fluid shear. Various crystallization phenomena can occur simultaneously in a crystallizer, and their importance may vary depending on local conditions. The growth rate may vary among crystal growth planes (faces), which can lead to needle-like crystals or other shapes. While hypothesis mechanisms provide expressions that describe the kinetics for most crystallization kinetic phenomena,[2] the parameters in these expressions must be estimated experimentally. However, the exact estimation of the parameters is impossible due to the limited and noisy data and the strong sensitivity of most crystallization processes to trace unmeasured chemical species in the feedstocks.

This motivates using model identification techniques that quantify the accuracy of the model parameters.[3,4]

A robust model identification procedure is developed for multidimensional crystallization processes. This procedure includes parameter estimation, model selection, and model-based optimal experimental design applied to infrared and particle size distribution measurements. A population balance model is used to model the multidimensional crystallization process. Nonlinear parameter estimation techniques are used to compute the kinetic parameters that best fit the data. Statistical techniques are used to calculate a confidence region that quantifies the accuracy of the model parameters. This confidence region is used by the optimal experimental design procedure to determine the best configuration and conditions for the next batch experiment. The confidence region is also used to select candidate models for the fundamental physical phenomena, and to assess the effectiveness of the best simulation model in representing and predicting the experimental observations.

The next section describes the model identification procedure. The experimental setup and apparatus are described in Section 3. Then, the approach is demonstrated for the crystallization of potassium dihydrogen phosphate from aqueous solution.

2. Model Identification

Model identification is an iterative procedure. Initial estimates of the kinetic parameters are used to design the first experiment, which is implemented in the laboratory. The collected experimental data are used to compute improved parameter estimates and an associated confidence region, which are used to design the next laboratory experiment.[4] A detailed study has shown that estimates obtained from this procedure can be more accurate than estimates obtained without optimal model-based experimental design.[5] The relationship among model selection, parameter estimation, and optimal experimental design is illustrated in Figure 1.

2.1. *Model development*

A population balance equation can account for the distribution in crystal dimensions, purity, and other state variables.[6] While the approach is rather general, for specificity the model equations will be written for the two-dimensional batch crystallization process in the laboratory. In this case the crystal shape is characterized by two characteristic lengths, r_1 and r_2.[7] Supersaturation is the driving force for nucleation and crystal growth. The relative supersaturation is

$$S = (C - C_{sat})/C_{sat} \tag{1}$$

where C is the concentration of solute, and C_{sat} is the saturation concentration. Let $f(r_1, r_2, t)$ be the crystal distribution function:

$$f(r_1, r_2, t)dr_1dr_2 = \text{the number of particles in the system in the range } r_1 \pm dr_1/2 \text{ and } r_2 \pm dr_2/2 \text{ at time } t \tag{2}$$

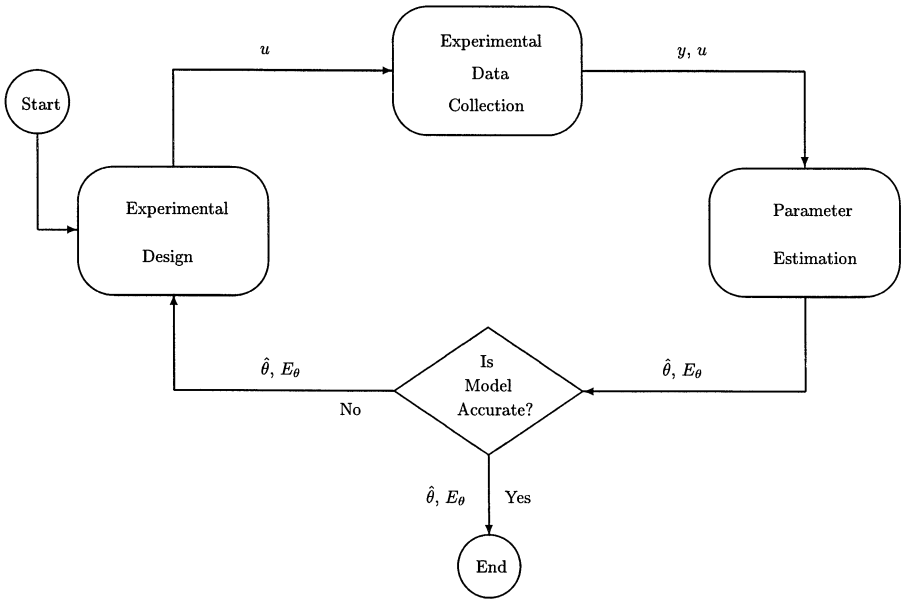

Fig. 1. Iterative process of model identification: u represents all experimental design variables (e.g., seed distribution, temperature profile), y represents the measurements (e.g., infrared spectra, crystal size distribution), θ is the vector of kinetic parameters, and E_θ is a confidence region for the parameters.

Assume that the volume of the batch crystallizer is constant, the crystal slurry is well-mixed, the growth rate along each axis is size-independent, nucleated crystals have negligible size, and there is no crystal breakage or agglomeration. Then the population balance equation for a batch crystallizer is

$$\frac{\partial f}{\partial t} + G_1 \frac{\partial f}{\partial r_1} + G_2 \frac{\partial f}{\partial r_2} = B(C,T)\delta(r_1)\delta(r_2) \tag{3}$$

where G_1 and G_2 are the growth rates of r_1 and r_2 respectively, B is the nucleation rate, and $\delta(\cdot)$ is the Dirac delta function.

While the model identification procedure is directly applicable to the population balance equation (3), the method of moments can be used to replace (3) with a small number of ordinary differential equations:[6]

$$\begin{aligned} \frac{d\mu_{00}}{dt} &= B \\ \frac{d\mu_{ij}}{dt} &= iG_1\mu_{(i-1)j} + jG_2\mu_{i(j-1)}, \qquad i+j>0 \end{aligned} \tag{4}$$

where the ij cross-moment is

$$\mu_{ij} \equiv \int_0^\infty \int_0^\infty r_1^i r_2^j f(r_1, r_2, t) dr_1 dr_2. \tag{5}$$

The crystallization kinetics are typically written in terms of the supersaturation. The most widely adopted models are in power law form:[2]

$$B = k_b S^b \mu_{21} \tag{6}$$

$$G_1 = k_{g1} S^{g_1} \tag{7}$$

$$G_2 = k_{g2} S^{g_2} \tag{8}$$

where S is the relative supersaturation defined in (1), and k_b, b, k_{g1}, g_1, k_{g2}, and g_2 are kinetic parameters. Three additional parameters, the activation energies for nucleation and the two growth axes, are required when there are significant temperature variations.

A solute mass balance completes the model of the crystallizer. The amount of solute leaving the solution must be accounted by crystal growth:

$$\frac{dC}{dt} = -2\alpha G_1 \mu_{11} - \alpha G_2 \mu_{20} \tag{9}$$

where α is the geometric-stoichiometric ratio chosen so that the right hand side of equation has the correct units. This equation makes the assumption that the geometric-stoichiometric ratio is constant throughout the experiment.

2.2. *Parameter estimation*

Let $\boldsymbol{\theta}$ be the vector of kinetic parameters,

$$\boldsymbol{\theta} = \begin{bmatrix} g_1 \\ k_{g1} \\ g_2 \\ k_{g2} \\ b \\ k_b \end{bmatrix}. \tag{10}$$

Then the estimation problem is to minimize

$$\Phi(\boldsymbol{\theta}) = \sum_{i=1}^{N_m} \sum_{j=1}^{N_{d_i}} w_{ij} (y_{ij} - \tilde{y}_{ij})^2 \tag{11}$$

where y_{ij} and $\tilde{y}_{ij}$ are the measurement and model prediction of the ith measured variable at the jth sampling instant, w_{ij} is a weighting factor, N_m is the number of measured variables, and N_{d_i} is the number of sampling instances for the ith measurement. The weights are selected based on maximum likelihood or by using estimates for the standard deviation of each measurement.[8]

Due to random errors in the measurements, the parameter estimates are stochastic variables with probability distributions. These distributions can be used to estimate the hyperellipsoidal confidence region that quantifies the accuracy of the parameters:

$$E_\theta \equiv \{\boldsymbol{\theta} : (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^T \mathbf{V}_\theta^{-1} (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}) \leq \chi^2_{N_p}(\alpha)\} \tag{12}$$

where α is the confidence level, N_p is the number of parameters, $\chi^2_{N_p}$ is the chi-squared distribution with N_p degrees of freedom, and the parameter covariance matrix $\mathbf{V}_\theta$ can be computed via linearization or Monte Carlo simulations.[8]

2.3. *Model selection*

There are a number of expressions for crystallization kinetics. Depending on the mechanistic details, alternatives for the nucleation kinetics (6) are

$$B = k_b S^b \mu_{21}^2 \tag{13}$$

and

$$B = k_b S^b \mu_{11} \tag{14}$$

which are two-dimensional generalizations of one-dimensional nucleation models.[2,9] While statistical methods for choosing different model structures are available, alternative models are commonly compared in terms of the quality-of-fit or the accuracy of the kinetic parameters.[4]

2.4. *Optimal model-based experimental design*

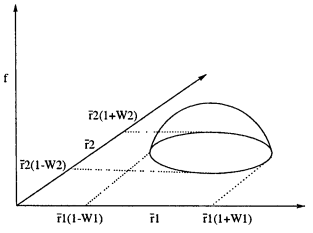

Fig. 2. Seed crystal size distribution: $\bar{r}_1$ and $\bar{r}_2$ are the mean characteristic lengths; W_1 and W_2 are the widths of the distribution.

Reduction of the uncertainty in the kinetic parameters is the primary goal of model-based experimental design, which computes the experimental conditions that minimize the volume of the hyperellipsoid (12). Sequential model-based experimental design takes into account all the data from past experiments while maximizing the informativeness of the data collected in the next experiment.[8] The objective is to minimize

$$\Psi(\mathbf{u}(t), \hat{\boldsymbol{\theta}}) = \det(\mathbf{V}_\theta) \tag{15}$$

where $\mathbf{V}_\theta$ is the covariance matrix based on both current and past data. In (15), $\mathbf{u}$ is the experimental design variables, which can be functions of time t. For the process considered in the next section, $\mathbf{u}$ represents the temperature and seed characteristics, where the seed is characterized by its initial mass m_{seed}, two mean

characteristic lengths $\bar{r}_1$ and $\bar{r}_2$, and the percentage widths W_1 and W_2 (see Figure 2). Then the optimization problem is to minimize (15) subject to the experimental constraints $\mathbf{u}(t) \in \Omega_{\text{exp}}$. For example, the temperature must stay within the operating range of the crystallizer. Constraints on the states can be included in Ω_{exp} by parameterizing the state variables in terms of $\mathbf{u}(t)$, then writing the requirements in terms of $\mathbf{u}(t)$.

3. Experimental Apparatus

The batch cooling crystallization apparatus has (1) a sampling port, (2) a thermocouple, (3) a Fourier Transform Infrared (FTIR) Spectrometer with Attenuated Total Reflection (ATR) probe, (4) a model M400L Lasentec Focused Beam Reflectance Measurement (FBRM) probe, and (5) a model 700L Lasentec Particle and Vision Measurement (PVM) system with fiber optic probe. The *in situ* FTIR-ATR probe provides multiple reflections along the length of a ZnSe crystal inserted in solution so as to magnify the infrared signal. The ZnSe crystal was designed so that a volume exclusion effect causes the measured infrared signal to be based almost entirely on the concentration in the solution phase, and not on the crystals in the slurry.[10,11] The construction of the calibration curve that relates the infrared spectra to the solution concentration for potassium dihydrogen phosphate(KDP)-water slurries is described elsewhere.[12] The FBRM measures the chord length distribution, which is a single distribution that is a weighted average of the multidimensional crystal size distribution. The PVM provides snapshots of the crystals in real time.

The temperature of the slurry is specified using a jacket, a valve that sets the ratio of hot and cold water flow rates to the jacket, and a proportional-integral control system designed via Internal Model Control.[13] A separate Nikon optical microscope is used to measure the characteristic lengths of crystal samples to determine the crystal size distribution.

4. Identifiability and Measurement Noise

Before carrying out experiments, it is useful to consider whether the measurements may have enough information to estimate all of the kinetic parameters (this problem is known as *identifiability*). Since the solution concentration provides no information on the relative growth rates associated with each crystal axis, the solution concentration by itself does not provide enough information to estimate the kinetics for multidimensional crystallization processes. Some straightforward analysis indicates that measurements of the solution concentration and the average length along each axis do provide enough information to estimate all the nucleation and growth kinetic parameters.

A second consideration is the relative magnitude of measurement errors associated with the various cross-moments of samples of slurry collected from the crystallizer. The very small crystals are unobservable in the optical microscope used to measure the size, so the measurement of μ_{00}, which is the total number of crystals

per mass of solvent, will be inaccurate. Also, μ_{ij} will be inaccurate for large values of $i+j$, since these cross-moments will be sensitive to the low-sampling statistics of the large particles. Therefore measurements of μ_{00} and μ_{ij} for large $i+j$ should not be used when estimating the crystallization kinetics.

5. Results and Discussion

Table 1 shows the results of applying the model identification procedure. The optimal seed mass for both experiments was to use 20 grams of KDP for the 2000 grams of solvent (this was the lower bound, which was specified by the accuracy of the solution concentration measurement). A large number of size ranges and distribution widths gave similarly informative data.

After the first experiment, model selection was performed for the three nucleation models (6), (13), and (14). The nucleation model (13) gave a very poor fit to the data, whereas (14) gave a very similar fit as (6). This agrees with Miller, who stated that it can be difficult to determine the precise nucleation mechanism based only on solution concentration and size distribution data.[14]

Table 1. Kinetic parameter estimates and their standard deviations using (6) as the nucleation model. The standard deviations were computed via Monte Carlo simulations.

Parameters	Experiment 1	Experiment 2	units
g_1	1.42±0.06	1.36±0.05	dimensionless
k_{g1}	431±199	451±141	μm/min
g_2	1.63±0.07	1.62±0.06	dimensionless
k_{g2}	2537±743	3308±757	μm/min
b	2.06±0.16	1.79±0.1	dimensionless
k_b	1.28E6±0.21E6	2.26E6±0.48E6	#particles/cm^{3}min

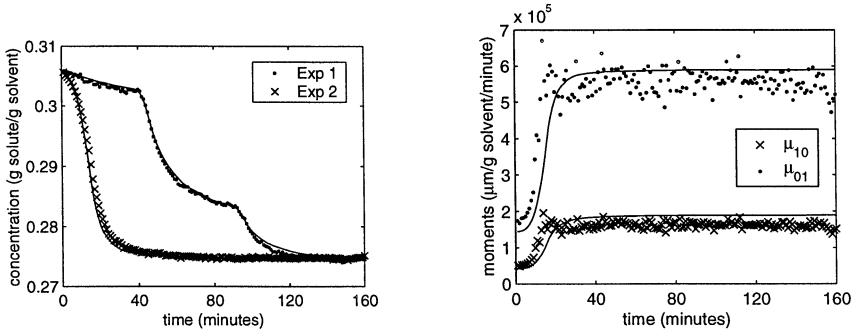

Fig. 3. The measured solution concentrations for both experiments and the measured moments (μ_{10} and μ_{01}) in experiment 2 along with the simulation profiles (solid lines).

The first experiment used samples collected for every 30 minutes to estimate

the cross-moments. The second experiment caused a large amount of nucleation to occur. Due to the very large number of small crystals relative to the number of seed crystals, it was impossible to obtain accurate estimates of the cross-moments via sampling, so a weighted normalization was used to estimate the moments from FBRM data collected for every minute.[15]

The growth parameters are quite consistent for the two experiments. Figure 3 verifies the proposed model identification procedure in a multidimensional crystallization process. While the nucleation parameters are significantly different, the parameter changes in the two experiments have a compensatory effect, so that the nucleation rates B are approximately equal for the supersaturation values obtained during most of the experiments. This correlation, which is not captured in standard deviations written for each parameter, can be quantified using the parameter covariance matrix (12).

Acknowledgments

Funding from Merck and DuPont is acknowledged.

References

1. T. Togkalidou, R. D. Braatz, B. K. Johnson, O. Davidson and A. Andrews, *AIChE J.* **47**, 160 (2001).
2. J. Nyvlt, O. Sohnel, M. Matuchova and M. Broul, *The Kinetics of Industrial Crystallization* (Elsevier, Amsterdam, 1985).
3. D. L. Ma, S. H. Chung and R. D. Braatz, *AIChE J.* **45**, 1469 (1999).
4. H. B. Matthews and J. B. Rawlings, *AIChE J.* **44**, 1119 (1998).
5. S. H. Chung, D. L. Ma and R. D. Braatz, *Chemo. Int. Lab. Syst.* **50**, 83 (2000).
6. A. Randolph and M. A. Larson, *Theory of Particulate Processes* (Academic Press, San Diego, 1988).
7. D. L. Ma and R. D. Braatz, *IEEE Trans. Cont. Sys. Tech.* **9**, 766 (2001).
8. Y. Bard, *Nonlinear Parameter Estimation* (Academic Press, New York, 1974).
9. E. P. K. Ottens, A. H. Janse and E. J. de Jong, *J. Crys. Growth* **13/14**, 500 (1972).
10. R. D. Braatz and S. Hasebe, in *Chemical Process Control VI*, ed. J. B. Rawlings and B. A. Ogunnaike (AIChE Press, New York, 2001), in press.
11. D. D. Dunuwila and K. A. Berglund, *J. Crys. Growth* **179**, 185 (1997).
12. T. Togkalidou, M. Fujiwara, S. Patel and R. D. Braatz, *J. Crys. Growth* **231**, 534 (2001).
13. R. D. Braatz, in *The Control Handbook*, ed. W. S. Levine (CRC Press, Boca Raton, Florida, 1995), p. 215.
14. S. M. Miller, Ph.D. Thesis, Univ. of Texas, Austin, 1993.
15. A. Tadayyon and S. Rohani, *Part. Part. Syst. Char.* **15**, 127 (1998).