Model Identification and Control of Batch Crystallization for an Industrial Chemical System

by

Hazel Benton Matthews III

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

(Chemical Engineering)

at the

UNIVERSITY OF WISCONSIN–MADISON

1997

To my parents, may I always be so lucky as to be so loved...

Acknowledgments

During my graduate studies I have been fortunate to experience academic life at two extraordinary universities. The University of Texas at Austin and the University of Wisconsin-Madison will each live long in my memory both as academic institutions and as social centers. Austin and Madison have been wonderful hosts and I hope to visit them often in the future. For those who have helped me along during my graduate tenure I give my humble thanks and my humble apology should I somehow forget you in the notes that follow.

I would like to thank the members of my academic committees both here at Wisconsin and also at Texas for taking the time to read my ramblings and for occasionally help me right my ship in time of need. At Wisconsin, I would like to thank Professors Warren Stewart, Ross Swaney, Douglas Cameron, and Richard Hartel. In particular, I would like to thank Professor Stewart for enlightening my graduate experience through his publications, consultations and kind words. At Texas, I owe thanks to Professors Thomas Edgar, David Himmelblau, and Roger Bonnecaze.

In both Madison and Austin, I was lucky to share the company of my advisor Professor James B. Rawlings. His timely comments and seemingly uncanny ability to solve in five minutes problems that I had wrestled with for five weeks, were invaluable to my growth as an engineer and scientist. Personally, it has been a pleasure spending time on the golf course, hockey rink and croquet course with Dr. Rawlings, his wife Cheryl and their children Melanie and Blake. I greatly enjoyed the many picnics and TGIF's that I shared with the Rawlings family.

Perhaps the greatest contribution that Professor Rawlings made to my graduate experience was to select and maintain a research group as full of hilarity as of intellect. It has been a fine time working with Team Rawlings. From my time as research child I was graced by the likes of Andrew Fordyce, John Eaton, Scott Meadows, Steve Miller, Ken Muske, and Tony Mullins. Fine fellows all, though I owe particular thanks to John, Steve and Ken. John was always there to set my compass to true North on any imaginable computer issue or to enliven my day with an ice cream run. His knowledge of latex, octave and fortran saved me more times than I care to count and I cannot thank him enough. Steve was my predecessor in crystallization and, along with fine fellowship, he left a thesis and a code that greatly facilitated my work and diminished my pain. Ken was a good friend and a fine fisherman and an improving skier. Looking forward to being in the boat or on the mountain with him again soon.

As I have in turn become the old man in the group, I have happily enjoyed the company of Rahul Bindlish, Chris Rao, and Scott Middlebrooks. Thanks Rahul for about a million good-hearted smiles and thank you Scott for working a little musical culture into my life. Chris, I must thank for being a constant source of amazement. His intellect, attitude, oddity and extremism have made for memories that I will not soon forget (e.g. "Both wore shoes; one wore a bike helmet."). John Campbell came into the group a year after I did and has been there all along. Nice sharing the ride with you, John, hard to believe its over.

Of the really new kids on the Rawlings block, I have been fortunate to know and enjoy the company of Daniel Patience, Hazim Mohameed, Jenny Wang, and Jenny Harting. In particular, I wish Daniel luck as he takes up the sword to slay the next crystallization beast. To our international contingent I say a heartfelt thanks for introducing us to your cultures and your studies. In this fine group are Pierre Scokaert, Peter and Rolf Findeisen, and Olav Slauphaug. Peter had the unfortunate task of sharing my office as this document came to a close, he was a good man to lean on in a time of need. I wish him luck as he heads back to Germany.

The crystallization project also allowed me to spend some time working with Eastman Kodak in Rochester, New York. I made many friends there but wish to thank in particular Mary Guinan, Michael Boroson, Peter Bessey, and my pseudo-boss and good buddy Mary Flohr. My visits would not have been possible, nor as enjoyable, without the technical guidance of Dr. John Hamer. Dr. Hamer was kind enough to be on my preliminary committee at The University of Texas and I owe him my fondest thanks.

At the University of Wisconsin, special thanks are due to Todd Ninman, Donna Gable and Roger Packard for making the move up from Texas survivable. Todd, in particular, deserves my thanks for many small favors and for the huge favor of keeping the VAX alive long enough for me to complete my experimental work. Al Hansen and John Cannon were also invaluable for their assistance and patience as I reconstructed my laboratory. Thanks to you both.

Lastly, I must thank the special friends who have made my life such a joy for the past few years. Thank you to Natalie Gould and Bridget McLain for being the friends from back home that helped make trips back home so special. Thank you to David Coie for moving to Austin and buying a fine home and always being Dave. It was a wonderful thing to have a true best friend and more wonderful still that he could introduce me to the love of my life, Sylvia Rosa Santaballa. There are no words adequate to thank her for all she has done to make my life worth celebrating. She is oasis and guidepost and inspiration. To my parents and sister I can only say that I am lucky without measure to have so fine a family. I cannot wait to come home.

Hazel Benton Matthews III

University of Wisconsin–Madison

April 1997

Contents

Model Identification and Control of Batch Crystallization for an Industrial Chemical System

Hazel Benton Matthews III

Under the supervision of Professor James B. Rawlings

At the University of Wisconsin–Madison

Batch crystallization is an important separation process used in the production of low volume, high value-added chemicals. Often, the final-time crystal size distribution (CSD) of a batch crystallization slurry determines product qualities such as appearance or slurry flow characteristics. The CSD also influences strongly the efficiency of solid-liquid separation processes such as filtration and drying. These downstream processes may be facilitated by controlling the evolution of the CSD during the crystallization step. This study examines the improvement of slurry filtration through the determination of optimal temperature profiles for seeded, batch cooling crystallizers. The temperature profiles are optimized subject to a fundamental model of the dynamics of the CSD. The model is identified for a photochemical-heptane system that is currently in production at the Eastman Kodak Company. This system exhibits some of the problems associated with industrial crystallizations including slow growth rates, an irregular crystal habit and a size distribution that is difficult to characterize.

The population balance equation provides a convenient way to model particulate phase processes like crystallization. Using a population balance structure, this study identifies the kinetic models describing crystal nucleation and growth of the photochemical system in a 3-liter seeded, batch cooling crystallizer. A nonlinear optimization code maximizing the posterior density function of the parameters with respect to the data is used to infer the kinetic parameters from on-line measurements of liquid phase solute concentration and slurry turbidity. New models are identified to account for crystal habit dynamics and size-dependent nucleation. Linear 95% confidence intervals are determined to summarize the uncertainty in the optimal parameter estimates. In order to optimize the information content in the experimental data sets, two of the experiments are designed using optimal experimental design techniques.

Given the identified model, the optimization of the process is considered. Generally, the filtration of a crystallization slurry is facilitated when the particles are large and the size variance of the population is small. However, because seeded batch crystallizers tend to produce bi-modal size distributions consisting of a class of seeds and a class of nucleated crystals, the goals of large size and small variance are difficult to reconcile. Therefore, this study focuses on the effect of minimizing the final-time mass of crystals in the nucleated class relative to the mass of the seed crystals (m_N/m_S). The objective is minimized with respect to a piecewise-linear temperature profile to be applied during the crystallization.

The optimal profile is determined using a nonlinear optimization that accounts for final-time and state constraints. This method guarantees that the solution adheres to realistic process limitations such as maximum cooling rates and yield constraints. The sensitivity of the optimal temperature profile to seed mass, run duration, and parameter uncertainty are analyzed. It is shown that increases in seed mass or run duration translate into improvements in the optimal value of the objective function.

Actual improvements in filtration resulting from implementation of two optimized input profiles are quantified experimentally by calculation of the average specific resistance of the filter cake. The filtration results for the optimal experiments are compared to filtration results from the model identification experiments. The slurries produced by the optimal profiles give the lowest resistance values recorded during the study and the total filtration times for the controlled runs are shorter despite higher solids densities. The optimal profile with larger seed load gives a cake resistance 25% lower than the best identification experiment.

List of Tables

List of Figures

Notation

Lower Case Letters

b — supersaturation order of nucleation

$\boldsymbol{b}$ — general linear constraints

c_f — mass of crystals/mass of filtrate

c_p — solution heat capacity

c_v — population coefficient of variation

c_{vw} — weight-mean coefficient of variation

d — supersaturation order of dissolution

d_w — depth-to-width ratio

d' — supersaturation order of crystal removal

$\boldsymbol{e}$ — prediction error matrix

$f(L,t)$ — crystal size distribution

$f_k(L,t)$ — crystal size distribution in the kth stream

f_l — linear seed size distribution intercept

$f_s(L)$ — seed size distribution

$\overline{f_s(L)}$ — seed size distribution on m_c/V basis

g — supersaturation order of growth

$\boldsymbol{g}$ — general inequality constraints

h — conversion from solvent mass to slurry volume

$\boldsymbol{h}$ — general equality constraints

j — third moment order of nucleation

k_a — area shape factor

k_b — nucleation rate constant

k_{bd} — bulk diffusion rate constant of two-step growth

k_{bcf} — BCF spacing dependent parameter

k_d — dissolution rate constant

k_d' — removal rate constant

k_f — disturbance model constant

k_g — growth rate constant

k_g' — BCF model diffusion rate constant

k_{sr} — surface reaction rate constant of two-step growth

k_v — volume shape factor

l — beam length of transmittance sensor

l_w — crystal length-to-width ratio

$\overline{l_w}(t)$ — seed distribution averaged crystal length-to-width ratio

$l_w(0)$ — initial crystal length-to-width ratio

$l_w(\infty)$ — crystal length-to-width ratio at infinite ΔL

$\ell(\boldsymbol{\theta}|\boldsymbol{Y})$ — likelihood function of the parameters given the data

q_f — mass flow rate of filtrate

m — number of measurement types

m_c — crystal mass

m_f — filtrate mass

m_N — mass of nucleated crystals

m_s — solvent mass

m_{sl} — slurry mass

m_S — mass of seed crystals

m_z — crystal mass above minimum size

p — number of model parameters

$p(\boldsymbol{\phi}, \boldsymbol{Y})$ — joint probability density of parameters and data

$p(\boldsymbol{\phi}|\boldsymbol{Y})$ — conditional probability density of parameters given data

$p(\boldsymbol{Y})$ — prior probability density of the data

$p(\boldsymbol{\theta})$ — prior probability density of the model parameters

$p(\boldsymbol{\sigma})$ — prior probability density of the prediction error standard deviations

$p(\boldsymbol{\theta}, \boldsymbol{\sigma})$ — joint probability of model parameters and prediction error standard deviations

q_d — mass flow rate in the densitometer stream

r — surface reaction rate supersaturation order

t — time

t_f — final time of an experiment

$\boldsymbol{v}$ — particle velocities

$\boldsymbol{v}_y$ — particle velocities along internal coordinates

w — mass of dry filter cake/filtration area

w_k — prediction error weight for measurements of the kth type

$\boldsymbol{u}(t)$ — general manipulated variable vector

$\boldsymbol{x}$ — external coordinates (spatial coordinates)

$\mathbf{x}(\mathbf{t})$ — general model states

y_i — ith measurement

$\tilde{y}_i$ — ith measurement prediction

$\boldsymbol{y}$ — internal coordinates of particles

$\boldsymbol{z}$ — general set of coordinates

Upper Case Letters

A — jacket water heating area

$\boldsymbol{A}$ — matrix of linear constraint constants

A_f — filtration area

$A_p(L)$ — size-dependent projected area of a particle

A_s — crystal surface area

B — crystal birth rate

B° — crystal nucleation rate at size L_0

$\boldsymbol{B}_{ik}$ — sensitivity of the ith response of the kth type

C — solution phase concentration (mass solute/volume solvent)

$\hat{C}$ — solution phase concentration (mass solute/mass solvent)

C_i — interface concentration for two-step growth

$\hat{C}_{\min}$ — minimum final-time concentration (yield constraint)

C_{sat} — saturation concentration (mass solute/volume solvent)

$\hat{C}_{\text{sat}}$ — saturation concentration (mass solute/mass solvent)

$\hat{C}_0$ — initial solution phase concentration (mass solute/mass solvent)

D — crystal death rate

E_b — activation energy of crystal nucleation

E_g — activation energy of crystal growth

G — crystal growth rate

$\boldsymbol{H}_\theta$ — Hessian of the model parameters

I — intensity of light transmitted through slurry

I_0 — intensity of incident light

$\boldsymbol{I}_n$ — Fisher's information matrix

L — characteristic crystal size

L_m — maximum seed crystal size

$L_{\min}$ — minimum crystal size for nucleation

L_0 — initial size of nuclei

$L(0)$ — initial crystal size

$\overline{L}$ — population mean crystal size

$\overline{L}_w$ — weight-mean crystal size

$\boldsymbol{L}$ — sensitivity matrix of $\boldsymbol{u}$ wrt $\boldsymbol{\theta}$

N_T — number of elements in $\boldsymbol{T}$

P — pressure applied at top of filter cake

P_l — hydraulic pressure across filter cake

P_m — pressure required to overcome filter medium resistance

$Q(L)$ — extinction efficiency factor

Q_k — volume flow rate in kth stream

$\boldsymbol{Q}_1$ — left singular matrix of $\boldsymbol{L}$

$\boldsymbol{Q}_2$ — right singular matrix of $\boldsymbol{L}$

R — ideal gas constant

R_m — resistance of filter medium

S — relative supersaturation

$S_{\max}$ — maximum supersaturation limit

T — slurry temperature

$\boldsymbol{T}$ — vector of vertices for piecewise-linear temperature profile

T_j — jacket water temperature

T_0 — initial slurry temperature

U — overall jacket heat transfer coefficient

V — slurry volume

$\boldsymbol{V}$ — measurement covariance matrix

$\boldsymbol{V}_\theta$ — parameter covariance matrix

$\boldsymbol{Y}$ — matrix of experimental data

$\tilde{\boldsymbol{Y}}$ — matrix of model predictions

Greek Letters

α — confidence level of χ_p^2

α_c — specific resistance of filter cake

$\overline{\alpha_c}$ — average specific resistance of filter cake

χ_p^2 — Chi-squared distribution with p degrees of freedom

δ — dirac delta function

$\boldsymbol{\delta u}$ — perturbation of $\boldsymbol{u}$

$\boldsymbol{\delta\theta}$ — perturbation of $\boldsymbol{\theta}$

$\boldsymbol{\delta\theta}_{\text{worst}}$ — worst-case perturbation of $\boldsymbol{\theta}$

ΔC — simple supersaturation

ΔH_c — heat of crystallization

ΔL — integral of crystal growth rate

ϵ — solids-free density

μ_f — filtrate viscosity

μ_i — ith moment of the CSD

$\mu_{n,i}$ — ith moment of the nucleated crystal class

$\mu_{s,i}$ — ith moment of the seed crystal class

ν_f — filtrate kinematic viscosity

Φ_{be} — Bayesian objective function

Φ_{c} — optimal control objective function

Φ_{d} — optimal experimental design objective function

Φ_{wls} — weighted least squares objective function

ρ — solution phase density

ρ_c — crystal density

ρ_f — filtrate density

$\mathbf{\Sigma}$ — matrix of singular values

$\boldsymbol{\theta}$ — vector of model parameters

$\hat{\boldsymbol{\theta}}$ — vector of optimal model parameters

$\boldsymbol{\sigma}$ — vector of measurement standard deviations

$\hat{\boldsymbol{\sigma}}$ — vector of modal measurement standard deviations

τ — mean residence time

τ_s — slurry turbidity

Caligraphic Letters

$\mathcal{L}$ — log-likelihood function

Chapter 1

Introduction

The importance of the crystallization process to the school-aged child praying for a snowy deliverance from unfinished homework is inherently clear. However, its importance to the chemical industry is often not as well understood. Crystallization is a technology used to isolate solid products from liquid phase melts and solutions. Its advantages over other separation processes such as extraction and distillation are its high product purities, low temperature environment and relatively low capital and operating costs. These qualities have made crystallization, and in particular batch crystallization, a mainstay process of the specialty chemical, photochemical, and pharmaceutical industries. In fact, Du Pont estimates that approximately 70% of their products pass through either a crystallization or precipitation stage [58]. At Kodak, where the photochemical properties of their products are especially intolerant of impurities, 90% of all chemical components are isolated via batch crystallization or precipitation [44].

Unfortunately, crystallization processes are not without problems. Product yield, purity, and crystal morphology may all pose challenges in a particular system. In addition, many operations have performance parameters that depend on the sizes or shapes of the product crystals. These parameters may concern the appearance of the final product, its dissolution rate in solvents or the human stomach, or the flow properties of the crystallizer slurry. Product constraints can usually be specified in terms of some characteristic qualities of the individual crystals (e.g. crystal size, shape, or purity). However, because individual crystals are not likely to all have the same size or purity, this suggests the existence of distributions of the characteristic qualities over the crystal population. Often, as with crystal size, the distributed characteristic not only constitutes a constraint on the quality of the product but also affects the dynamic evolution of the crystallization process itself. This leads to a complex interplay between the distribution of the characteristic over the crystal population, the macroscopic states of the crystallizer environment (e.g. temperature, liquid phase composition), and the desired final-time state of the system.

In order to improve the performance of industrial crystallizers, it is important to be able to describe how the characteristic qualities influencing a particular process are distributed over the crystal population. One approach to this problem is the population balance equation (PBE) developed in the 1960's by Rudd [132] and Hulbert and Katz [71]. The population balance is a conservation equation describing the movement of the crystal population through a state space comprised of the spatial coordinates and the characteristic variables. The utility of the PBE approach for solving problems in crystallization is elucidated in the text of Randolph and Larson [123]. These authors focus on the crystal size distribution (CSD) as the key characteristic of a crystallization process. Mass transfer rate, slurry filtration characteristics, and physical appearance are all strong functions of the CSD. One goal of this study is to show that the PBE approach may be used to model the CSD of a relevant industrial crystallization system.

The PBE-based crystallization model is not complete without the specification of the kinetic processes that create, destroy and convect the crystal population along the size coordinate. Generally, the most important of these processes are crystal nucleation and growth, though in some systems agglomeration, attrition and even dissolution models are required. The nucleation rate is the rate at which small crystals enter the population and the crystal growth rate is the rate at which existing crystals increase in size. These processes and their effects on the crystal population are intricately related. For instance, the transfer of mass into the crystalline phase is a function of available crystal surface area and the growth rate. However, the contributions of nucleation cannot be discounted because in many systems most of the available area is on the surface of crystals nucleated during the crystallization. Each process is driven by the nonequilibrium state of the system. In solution crystallization this state is approximated by the relative supersaturation, which is a measure of the difference between the liquid-phase composition of the crystallizing substance and its equilibrium composition.

In order to model a specific crystallization system, the parameters that relate growth and nucleation to supersaturation must be estimated from experimental data. Therefore, sensors capable of measuring the characteristics of CSD as they evolve during a crystallization are crucial to a model-based crystallization study. This study utilizes on-line measurements of slurry turbidity and liquid phase solute concentration. In essence, these are measurements of the total mass and total surface area of the crystals. Ideally, the CSD would be measured directly. However, none of the sensor technologies capable of such measurements in limited systems (light scattering, optimal microscopy, Coulter counters) are yet applicable for on-line measurements in more challenging industrial systems characterized by dense slurries and irregular crystal habits.

The information content in the data is enhanced by studying a batch crystallizer system. Batch systems are convenient both in the laboratory and on the production line. In industry they are generally used to isolate low volume, high value-added products (e.g. photochemicals and pharmaceuticals) and are valued for their flexibility. A particular batch reactor-crystallizer system may often be used to produce many different low volume products. In the laboratory, the dynamic nature of a batch crystallizer allows the collection of data over a wide supersaturation range without consuming a large amount of solvent or raw material stock. In this study, supersaturation is controlled by manipulating the temperature of the crystallizer, which changes the supersaturation by raising or lowering the solubility of the solute in the liquid phase. Other ways of controlling supersaturation include the addition or evaporation of solvents, chemical reactions, and manipulation of the feed rate in continuous crystallizers.

Miller and Rawlings [94] show that batch measurements of turbidity and concentration are sufficient to obtain precise estimates of kinetic parameters for the KNO_3-H_2O system using a maximum likelihood estimation method. However, no effort has been made to estimate parameters in more problematic industrial systems. This study uses Bayesian inference techniques to determine precise kinetic parameters for a photochemical-heptane system in production at Kodak. In order to maximize the information content in the data, optimal experimental design techniques are also applied.

The complex behavior of crystallization processes leads to significant control problems. Often continuous crystallizers are plagued by instabilities that produce cycling behavior in the product CSD. This behavior is a function of the supersaturation dependence of the growth and nucleation processes and is especially common in crystallizers utilizing classified product removal [138]. However, despite the interest such instabilities have spawned in the control of continuous crystallizers, there have been relatively few attempts to examine control policies for batch systems. Of the studies conducted, all but one [147] have focused on aqueous inorganic systems and none have evaluated the filtration performance of the optimized policy. This study uses the identified model and the tools of optimal control theory to develop and test two control profiles designed to improve the filtration qualities of the photochemical-heptane system. Constant pressure slurry filtration tests are used to verify the improvements garnered from the optimized control profiles.

The photochemical material examined in this study was provided by the Eastman Kodak Company and is presently produced at Kodak's plant in Rochester, New York. As such, its chemical name and structure are protected by a nondisclosure agreement. The primary motivation for choosing this system was to establish the utility of fundamental modeling, estimation, and control techniques for improving the performance of an industrially relevant crystallization. The photochemical system is by no means a model system. Crystal shape is irregular and may change during a crystallization, the solubility curve is only moderately well-determined, growth rates are very slow and are influenced by impurity levels in the raw material stock. These difficulties are not desirable but they are prevalent in industrial batch crystallization systems that, thus far, have not received attention from academic researchers. The lack of research performed on problematic systems is unfortunate given that organic systems of this type comprise a significant portion of the industrial applications of batch crystallization. In keeping with Kodak's industrial practice, the batch experiments conducted in this study are seeded crystallizations.

1.1 Thesis Overview

This thesis is divided into four main chapters that describe the experimental apparatus, the general model and solution techniques, the model identification study, and the determination and results of an optimal control study. These chapters are briefly summarized below.

Chapter 2 describes the batch cooling crystallizer used to collect experimental data for this study. The apparatus is a two-liter, jacketed glass crystallizer. On-line process measurements are provided by an *in situ* transmittance sensor (turbidity) and an external densitometer (concentration). Viable alternatives to these measurement devices are discussed as is a separate constant pressure filter apparatus that was used to filter the product slurries. Average specific cake resistance is the measure used to evaluate the filtration performance of the crystallization slurries. This measure is derived in Chapter 2. Typical operating procedures for the crystallizer and the filtration apparatus are also summarized in this chapter.

Chapter 3 describes the formulation of the general PBE and its subsequent simplification for CSD modeling in a batch crystallization system. Corresponding mass and energy balances are also presented, though their derivations are relegated to Appendix D. This study uses empirical power law models to describe the dependence of the crystal nucleation and growth processes on the relative supersaturation driving force. The secondary (collisional) nature of the nucleation process is taken into account by augmenting the supersaturation power law with an additional crystal mass term. A literature review of crystal growth and nucleation modeling is also provided.

In addition, Chapter 3 presents the moment formulation that is used to solve the population balance partial differential equation. Alternative solution methods are discussed. The chapter closes with a series of simulations solving the general nonlinear, dynamic moment model for a nominal set of parameters.

Chapter 4 describes the model identification phase of the study. The focus is on the use of experimental transmittance and concentration data to infer the system specific parameters of the nucleation and growth models proposed in Chapter 3. A review of kinetic model identification efforts in the crystallization field is presented as is the derivation of a multi-response, Bayesian objective function. The kinetic parameters are estimated via unconstrained nonlinear optimization of the Bayesian objective using the sequential quadratic programming code NPSOL [55]. The chapter follows the progression of the identification study through the initial experimental designs and data collection, to the identification of a bias in the data, to the removal of the bias and subsequent revision of the model, and finally to the design of an optimal experiment maximizing the information content of the data. Precisely determined kinetic parameter estimates selecting the mode of the Bayesian posterior density function are presented and the uncertainty of the optimal estimates is assessed by the calculation of linear 95% confidence intervals. Analyses of the prediction error and linearity assumptions are presented along with some final comments on alternate models and data sets.

Chapter 5 describes the optimal open-loop control of the photochemical-heptane system using the identified model of Chapter 4. A constrained nonlinear optimization is used to determine crystallizer temperature profiles that either maximize seed size or minimize the ratio of nucleated crystal mass to the mass of grown seeds ($m_N(t_f)/m_S(t_f)$). The sensitivity of the $m_N(t_f)/m_S(t_f)$-optimal profile to parameter uncertainty, seed mass and experiment duration is analyzed. The effect of the difference in the supersaturation orders of nucleation and growth on the optimal profile is also examined. Two $m_N(t_f)/m_S(t_f)$-optimal strategies (optimized for different seed masses and run durations) are implemented on the experimental apparatus and the specific cake resistance values from these runs are compared to corresponding measurements from the identification experiments. The optimized runs produced the lowest cake resistance values measured during the study. A change in the system kinetics that affected the optimal control experiments is also discussed.

Chapter 6 summarizes the identification and control results of the study and presents some ideas for future work.

Chapter 2

Experimental Apparatus

This chapter describes the experimental apparatus, measurement technologies, and chemical system used in this study. With regard to the measurement devices, special efforts are made to examine alternative technologies that are available for recording similar types of data and to explore the possible options for augmenting the existing apparatus with additional measurement types. As the bench-scale batch crystallizer apparatus is complemented by a constant pressure filtration unit, that apparatus is also described. Lastly, a step-by-step description of the crystallization and filtration procedures provides the reader with a sense of exactly how the experiments were conducted.

2.1 Apparatus Overview

The complexity of the mathematical models used to describe the crystallization process is contrasted by the simplicity of the physical apparatus itself. The bench-scale batch cooling crystallizer used in this study is little more than an agitated glass jacketed vessel. Even with the addition of external measurement devices and cooling water controls, the appara-

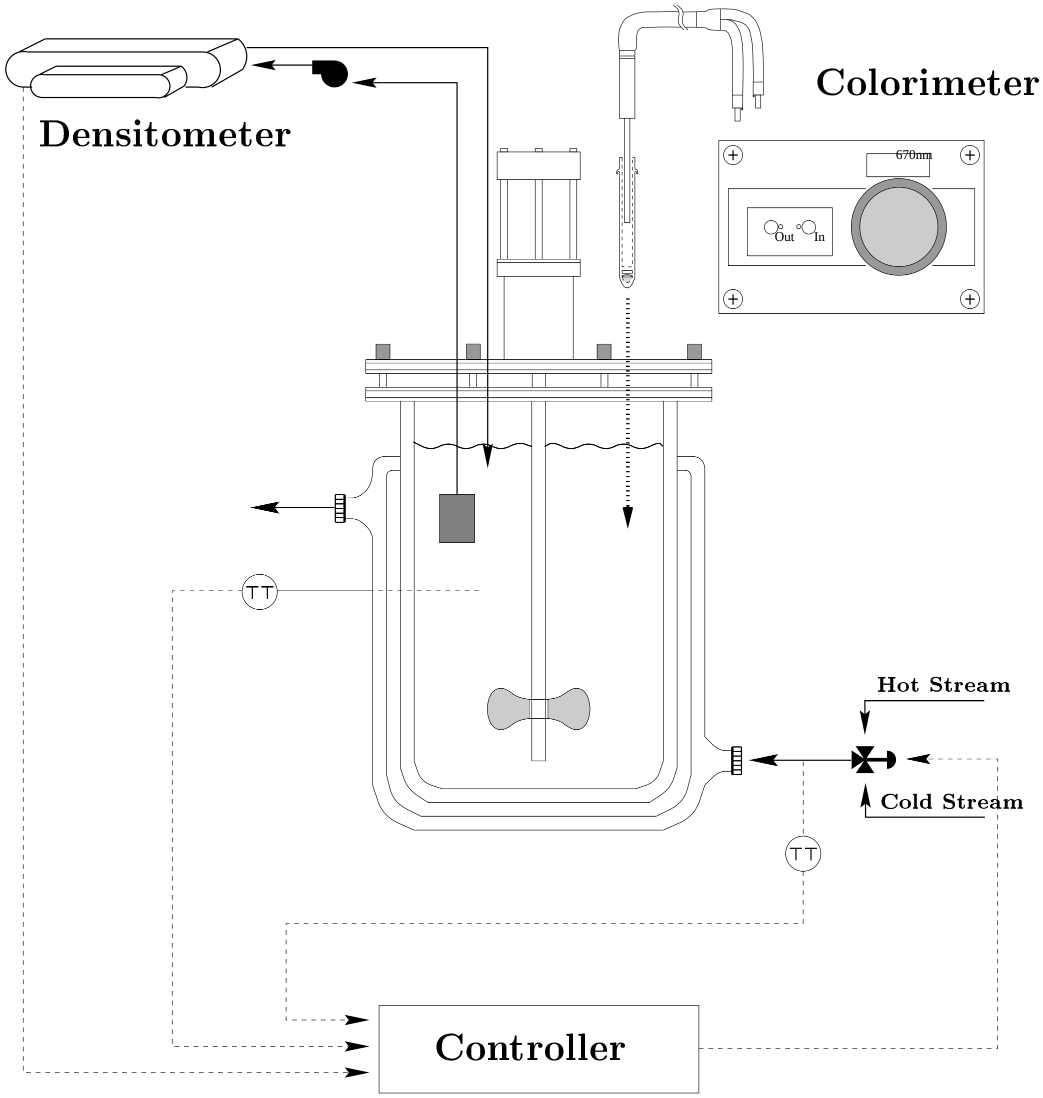

Figure 2.1: Bench-scale batch crystallizer.

tus schematic shown in Figure 2.1 remains quite modest. The apparatus is comprised of the crystallizer, a densitometer for on-line concentration measurement, an *in situ* probe for on-line turbidity measurement, and a model-based temperature control system. Additional elements of the system, implicit in the diagram, include data acquisition and control hardware and a workstation that performs on-line control calculations and stores the collected data.

The crystallization apparatus is designed to allow satisfactory determination of kinetic parameters for a variety of chemical systems. Witkowski [158] and Miller [95] used earlier iterations of this apparatus to successfully complete kinetic identification studies for the Napthalene–Toluene and Potassium Nitrate–Water systems, respectively. This work undertakes the study of an industrial photochemical that is currently produced by the Eastman Kodak Company. This system is chosen to explore the application of detailed mathematical modeling to the problematic crystallizations that often occur in industry. The goal is to determine to what extent standard particulate system modeling techniques may be used to describe the behavior of such systems and, having done so, to what degree their performance may be optimized.

2.2 Chemical System

The photochemical-heptane system used in this study is a bench-scale, high-purity version of a crystallization that Kodak presently performs in its production facilities in Rochester, New York.[1] This system was chosen because it is an economically important system that exhibits some common features that make modeling industrial crystallization systems a challenging endeavor. A brief list of the particular features that apply to this system is given below.

- The material is polymorphic.
- Solubility versus temperature data for the crystallized morphology were not available prior to this study.

[1]The name and chemical formula of the photochemical are protected by a nondisclosure agreement.

- The crystal habit is approximately parallel-piped but inconsistent and needle-like in some instances.
- The crystal size is relatively small ($< 100\mu$m seed crystals).
- The crystallization kinetics are influenced by trace impurities.
- The crystal size distribution cannot be accurately measured due to shape and size concerns.
- The crystallization kinetics vary significantly from batch-to-batch.
- The crystal growth rate is very slow.

These problems are partly offset by the system's wide metastable zone, which permits seeding at high levels of supersaturation. The large metastable zone allows greater flexibility in experimental designs, eliminates concerns about the dissolution of seeds, and minimizes the effect of the uncertainty in the solubility curve.

This study approaches these pathological features in different ways. Polymorphism and impurity concerns are handled by designing identification experiments that eliminate their effects. Fortunately, the effects of polymorphism are eliminated by the fact that only one of the morphologies has been crystallized in the laboratory. Purity is treated by using raw material from a single industrial batch. The problem of determining the size distribution of the seed crystals is treated by assuming a distribution shape and determining the parameters of that shape from model equations and microscopic measurements. Special modeling treatments for these problems are presented in detail in Chapters 3 and 4.

The inability to quantify the crystal size distribution is the greatest challenge presented by this system. Light scattering methods fail because the needle-like habit and small size of the crystals lead to violations of scattering theory assumptions. Miller [95] provides a discussion of such shape dependent errors in the CSD measurements of a particular light scattering device. The habit and size also prevent the material from being effectively segregated with mechanical sieves. Microscopy methods, scanning electron and optical, are useful and can provide qualitative measurements (habit, crystal quality, maximum size) regarding the photochemical CSD. However, due to sampling concerns, they are not suitable for quantitative evaluation of size distributions.

The inability to measure the CSD prevents the validation of the identified model by comparison to a measured value. This unfortunate condition is mitigated by the use of transmittance and concentration data to confirm that the model describes accurately the second and third moments of the CSD. A final qualitative model validation is provided by measuring the filterability of the crystallized slurries. This measurement, described in Section 2.6, is the only available confirmation that optimal operating policies based on the identified model are improving performance. Note, that the optimization criterion must be cast in terms of the moments of the CSD and, without the ability to measure the CSD, success must be assessed without direct confirmation of the qualities that the modeling effort sought to optimize. Improved filtration performance offers a powerful, if qualitative, validation of the predictive capacity of the identified model.

The raw material stock used in this study was dried final product collected from Kodak's production plant in Rochester. The data presented in this study were all collected using material from a single batch.

2.3 Crystallizer Assembly

The heart of the experimental apparatus is the crystallizer, pictured in Figure 2.2. The model used in this study is a flat-bottomed, jacketed, glass vessel produced by Applikon. The full volume is 3 liters but the actual process fluid volume seldom exceeds 2 liters in an attempt to preserve solute and solvent materials. The volume of the crystallizer jacket is approximately 1.2 liters. When crystallization data is being collected, the jacket is insulated with an inch thick blanket of foam rubber to dampen the effects of disturbances in the ambient temperature. The crystallizer and all other elements of the apparatus that contact slurry or crystal-free liquor are contained in adjacent bench scale fume hoods.

The glass vessel is fitted with a stainless steel headplate that is secured to the glass section by the bolts of a metal collar that sits below a glass lip at the top of the crystallizer. This arrangement allows the formation of an airtight seal to prevent the evaporation of crystallization solvent during an experiment. The only ports for the system are those in the headplate. Six millimeter stainless steel tubes pass through the headplate to remove and return the densitometer's filtered sample stream. If desired, similar ports permit continuous slurry samples to be sent to an external Malvern particle sizer. Special fittings allow the depth of these tubes to be adjusted as desired. A larger eighteen millimeter port is fitted with an *in situ* colorimeter probe. This probe measures the turbidity of the slurry. Other eighteen millimeter ports are available for use in the addition of seed crystals or to provide access should the densitometer filter need to be cleared manually.

A sealed impeller shaft is integrated into the headplate. The impeller is a marine type prop set to drive the process fluid in a downward direction. A removable, direct

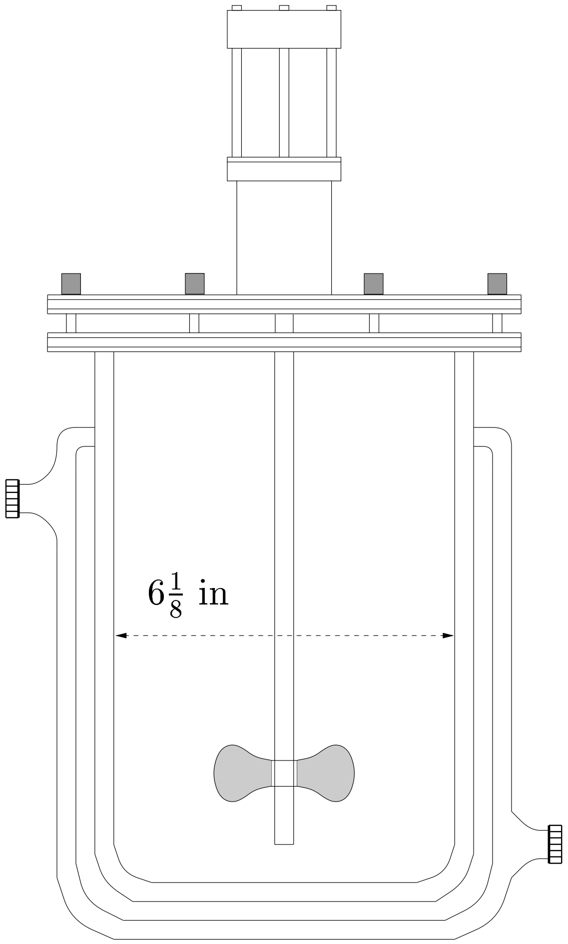

Figure 2.2: Bench-scale batch crystallizer.

current impeller motor with a speed range of 0 to 1250 revolutions per minute (rpm) is controlled by an external motor controller (Applikon, model ADI 1012). The high aspect ratio photochemical crystals used in this study are maintained in a well-mixed slurry for impeller speeds above 300 rpm. Unless otherwise noted, impeller speed is set to 350 rpm.

2.4 Control of Crystallizer Temperature

A batch cooling crystallizer apparatus must have the capacity to follow accurately and precisely a given temperature profile. This precise setpoint tracking is required to control the temperature-dependent relative supersaturation, S, that drives the rates of crystal growth and nucleation.

$$S = \frac{C - C_{\text{sat}}(T)}{C_{\text{sat}}(T)} \tag{2.1}$$

Where S is relative supersaturation, C is solute concentration, and $C_{\text{sat}}(T)$ is the temperature dependent saturation concentration, or solubility, of the solute. Crystallizer temperature is the sole control variable investigated in this study.

The temperature control system uses a hot stream/cold stream design to allow fast changes in jacket water temperature. The two streams are blended together in a three-way control valve to attain a constant flow-rate stream of the desired temperature. This design permits jacket water temperature settings anywhere within a range bounded by the temperatures of the hot and cold streams. Jacket temperature set points for the system are determined by a model predictive controller that regulates the temperature of the crystallization slurry. The specifics of the control system are addressed below.

2.4.1 Physical Subsystem

Jacket water for the crystallizer is drawn from the tap, filtered to remove particulate matter, and passed through a gear pump (Micropump, model 220/56C). The pump assures a constant flow rate through the jacket. The pump rate may be controlled by adjusting the output of a direct current drive (Warner Control Technologies, Bronco II model 169). For the experiments in this study the flow is maintained at a constant rate of approximately 2.3 liters per minute, corresponding to a jacket residence time of about 30 seconds. This assures that the jacket dynamics are fast relative to those of the crystallizer. The time constant for temperature changes in the crystallizer is approximately 8 minutes. After passing through the pump the flow is split into the aforementioned hot-side and cold-side streams.

Water entering the hot-side stream is sent through a series of two double-pipe heat exchangers. The heat exchangers are identical and together constitute a total heated length of approximately 2.1 meters. The jacket water stream runs through interior pipes of 1/4 inch diameter straight copper tubing. The annular regions contain heated water from a 9-kilowatt heater (Advantage Engineering, model S-915 21C1). Annular flow and temperature are maintained by the heater's 0.56-kilowatt pump and Sentra Temperature Controller, respectively. The flows are counter-current with annular water cycling from the heater to the second exchanger, to the first exchanger, and then back to the heater. The heater has a temperature range of 30°C to 180°C. However, except for the dissolution phases of some experiments, temperature was maintained at or below 45°C. An annular water temperature of 60°C provides sufficient energy to allow a crystallizer operating temperature of approximately 40°C. Generally, annular temperature is kept constant during the data collection phase of a crystallization experiment to minimize temperature disturbances. After leaving the heat exchangers, the hot-side stream is ported to one of two inlets of a 3-way control valve (Research). This valve blends the stream with the incoming cold-side stream to attain the desired jacket temperature. The valve has a "G" trim size, corresponding to a valve flow coefficient of 0.2, and is equipped with a 115 VAC EVA-1 actuator (Badger Meter Inc., model XB1.6A).

The flow split into the cold-side stream follows a path similar to that of the hot-side. This stream passes through approximately 14 meters of coiled 1/4 inch diameter copper tubing in a series of two insulated baths. These baths are filled with ice water as necessary. For crystallizer operating temperatures above 18°C, the tap water does not need to be cooled (Madison tap water temperature is approximately 14°C during the Summer months and is substantially colder in Winter). A short length of 1/4 inch Teflon tubing may be used to bypass the baths in such cases. After leaving the baths, the cooled stream enters the second of the two inlet ports in the 3-way control valve described above. The temperature of the combined stream is measured by 3-wire, 100Ω, platinum resistance temperature detectors (RTD's) (Omega Engineering Inc., model PR-13-1-100-1/8) at the jacket inlet and outlet ports. Similar RTD's provide on-line measurements of crystallizer, ambient, and densitometer return stream temperatures. After passing through the crystallizer jacket, the flow is sent to drain.

2.4.2 Controller Subsystem

The objective of the controller subsystem is to control the temperature of the contents of the crystallizer. This control is accomplished through a cascade design. A model predictive controller (finite horizon, quadratic dynamic matrix algorithm [47]) compares an on-line RTD measurement of the crystallizer temperature to a pre-programmed set point temperature and then calculates an appropriate corrective value for the temperature of the incoming jacket water. This value is passed, as a new set point, to a proportional-integral-derivative (PID) slave controller (Omega Engineering Inc., model CN-2002) that adjusts the 3-way valve to blend the hot and cold-side streams to meet the new goal. New jacket water set points are scheduled at a rate of one per minute to allow sufficient time for the model predictive controller to complete its calculations. A detailed description of the QDMC algorithm used by the model predictive controller is given by Miller [95].

The data acquisition and communications required to implement the controller design are presented in Section 2.7.

2.5 State Measurements for Crystallization Processes

For a cooling crystallization process the states of interest are the CSD, the concentration of the crystallizing solute in the liquid phase, and the system temperature. The CSD describes the state of the solid phase and its final-time value affects significantly the filtration qualities of the slurry. Concentration and temperature are required to specify the crystallization driving force in the system, recall Equation 2.1. As outlined in Section 2.2, it is not possible to measure explicitly the CSD for the photochemical-heptane system. In lieu of that measurement, this study measures the second and third moments of the CSD via transmittance and concentration measurements, respectively. Note the dual role of the concentration measurement.

This section describes the techniques used to perform the temperature, concentration and transmittance measurements in this study.

2.5.1 Temperature Measurement

Resistance temperature detectors (RTD) provide on-line temperature measurements of the crystallizer slurry, the jacket water inlet and outlet streams, the densitometer return stream, and the hood ambient. The data acquisition system records temperature measurements at 15 second intervals throughout a crystallization run. The time delay for the temperature measurement is negligible.

2.5.2 Concentration Measurement

In the literature several technologies have been used to measure solute concentration in crystallization systems with varying degrees of success. This section presents a detailed description of the densitometry method utilized in this study followed by brief reviews of three additional techniques.

Densitometry

First proposed by Garside and Mullin [50] for the on-line measurement of solute concentration, densitometry has been applied successfully for a wide array of crystallization systems in both integrated on-line [129, 127] and intermittent off-line [111, 112] sampling modes. Densitometry functions by correlating changes in the density of a solution to changes in the concentration of a solute in that solution. The measurement is typically made with a u-tube densitometer or an electro-magnetically balanced float.

This study used a u-tube densitometer (Anton Paar, model DPR 412 YE). This model has a 2 millimeter inside-diameter flow-through stainless steel u-tube and is equipped with a Pt100 RTD that measures the temperature of the solution as it passes through the cell. Temperature and u-tube oscillation period values are passed to a serial interface (Anton Paar, model DPR-S) that uses an RS-232C interface to communicate with the data acquisition computer.[2] A 25 milliliter/minute constant flow of solids-free solution is pumped through the tube by a peristaltic pump (Masterflex, model 7523-00). Crystals are prevented from entering the stream by a 15 micron filter at the inlet port inside the crystallizer. After

[2] More precisely, the counterweighted u-tube is excited electro-mechanically at a standard frequency which is altered by the density of the sample fluid within the tube. The actual measurement is a quartz pulse count that is inversely proportional to the vibrational frequency of the tube. The pulse count, or T-value, is the number returned by the instrument.

passing through the device the sample stream is returned to the crystallizer.

The densitometer measures the oscillation period of the u-tube and the temperature of the cell. For density measurement, oscillation periods and temperatures are recorded for solutions of known density. This calibration provides a reference table from which the density of new solutions may be determined by interpolation. Given a one-to-one relationship between solution density and solute concentration, concentration measurements are inferred similarly from on-line oscillation period and temperature measurements and a look-up table composed of the periods and temperatures corresponding to calibration solutions of known concentration.

Previous researchers that have used this experimental apparatus to examine Naphthalene/Toluene [158] and KNO_3/Water [95] systems have shown linear relationships between oscillation period and temperature at a given concentration and between oscillation period and concentration at a given temperature. These results were confirmed for the photochemical/heptane system as illustrated in Figure 2.3 by a typical calibration data set. Due to occasional drifts in the instrument readings, new calibrations were made as required.

Given the look-up table, a four-point interpolation scheme determines the solution concentration. A collected data point consists of an oscillation period and a solution temperature. The interpolation proceeds by first interpolating on the temperature scale to obtain a set of oscillation period/concentration pairs. Each pair corresponds to the temperature value of the data point on a particular constant concentration calibration set. Then a linear regression is performed on these pairs and the concentration is determined by the oscillation period of the data and the regressed parameters.

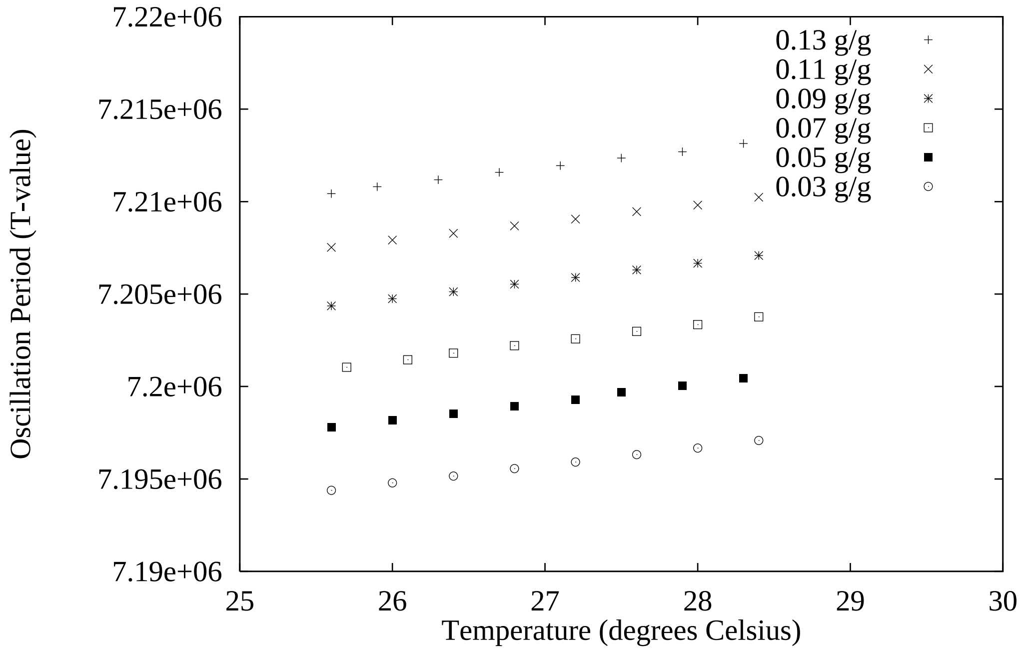

Figure 2.3: Densitometer calibration data collected from four reference solutions of known concentration.

The oscillation period measured by the densitometer is sensitive to temperature changes in the flowing solution. Such disturbances are minimized by maintaining a thermal equilibrium in both the instrument and its inlet stream. A 40 inch double-pipe heat exchanger is positioned between the non-isothermal crystallizer and the densitometer. The process fluid flows through 1/8 inch diameter aluminum tubing enclosed within an 3/8 inch stainless steel outer pipe. Water, recirculated and maintained at a constant 27.0°C by a water bath (Haake, model F4-C), flows counter-currently from the densitometer's own jacket into the outlet end of the heat-exchanger. By maintaining the measured solution temperature several degrees above the $18-23$°C operating range of the crystallizer, this arrangement is sufficient to damp a 2°C change in crystallizer temperature to approximately 0.1°C at the densitometer. To minimize the effects of ambient temperature disturbances, the densitometer is covered with 1/2 inch thick insulation. However, because such disturbances may still be significant, the ambient temperature of the hood containing the densitometer should be maintained at the same temperature at which the calibration data were collected. In this study, calibrations and data collection were performed with the hood maintained at 25.0°C. Ideally, as in to the work of Riebel et al. [129], the densitometer should be isolated in a temperature controlled box to eliminate all ambient disturbances.

After measurement in the densitometer, the process stream must be returned to the crystallizer at the appropriate temperature. The temperature adjustment is made by cooling the stream in a second 40 inch double-pipe heat exchanger. This exchanger has a 1/4 inch teflon internal tube enclosed by a 1/2 inch teflon external tube. Cooling water is recirculated by a second Haake bath and maintained at a temperature equal to the current set point of the crystallizer slurry minus 1.1°C. This setting is sufficient to cool the solids-free sample stream to the process temperature when the crystallizer is operated at approximately 20°C. Small mismatches in return temperature are tolerable due to the well mixed conditions in the crystallizer and the low mass flow-rate of the stream (approximately 17.5 grams/minute) relative to the total mass of the crystallization slurry (exceeding 1500 grams).

It was observed that the drag force of the inlet flow to the densitometer sample line resulted in the deposition of crystals on the face of the solids filter. To remedy this condition, frequent back-flushing of the filter was required. This was generally performed once every ten minutes to keep the filter clear and prevent removal of significant crystal mass from the bulk slurry. Under some conditions, this deposition effect changed the evolution of the growing crystal population to a significant degree. In these cases, replicate crystallizations were performed with intermittent operation of the densitometer loop. Typical operation would be 5 minutes of data collected every twenty minutes over the course of a 500 minute run. In the worst cases, particularly during periods of elevated nuclei production, crystals encrusting the filter were removed by sweeping the face of the filter with a rubber coated steel rod.

The method outlined above permits reproducible and accurate measurements of concentration to within $\pm 3.0 \times 10^{-4}$ grams photochemical/gram heptane. As with temperature data, oscillation periods and solution temperatures from the densitometer are transmitted on-line to the VAX computer via the data acquisition hardware and are given appropriate time stamps. In on-line operation, data is collected at intervals of approximately 10 seconds when the densitometer sample line is running. The dead time of the measurement due to flow to the instrument and device processing time is approximately one minute.

Alternate Technologies

Like densitometry, the other methods of concentration measurement that have been applied in crystallization systems measure changes in some concentration dependent physical characteristic of the continuous phase. Subsequently, implementation of these techniques also involves the generation of a detailed look-up table of calibrated values from solutions of known concentration. The three techniques summarized below include electrical conductivity, refractive index, and attenuated total reflection (ATR) infrared spectroscopy. Apparent advantages and disadvantages of each technique are outlined and successful implementations from the literature are referenced. In general, these techniques are not as precise as densitometry but do have the advantage of being *in situ* measurements that do not require separation of the solid phase prior to measurement.

Electrical conductivity measurements have been used by a small body of researchers in the crystallization field [103, 64, 88]. The primary advantage of the conductivity measurement technique is its insensitivity to the presence of the particulate phase. However, this feature is accompanied by several disadvantages: the technique is sensitive to temperature disturbances, the signal-to-noise ratio may be low for highly conductive systems (e.g. inorganic salts in water), there may be a conductivity maximum with respect to solute concentration in some systems [103]. Hložny et al. [64] provide an example of a successful implementation of the technique. These authors describe the use of conductivity to measure concentration in an aqueous ammonium aluminum sulfate system. They recorded approximately 30 on-line measurements over the course of a four hour batch cooling crystallization and report a measurement accuracy of $\pm 1 \times 10^{-3}$ gram solute per gram water.

Refractometry infers concentration changes from changes in the refractive index of a solution. This technique was used by several authors in the 1970's [98, 139, 63]. In a recent comparative study of refractometry and conductivity, Nývlt and coworkers [103] perform the measurement by transmitting light through a glass u-tube in contact with the crystallization slurry. A portion of the light reaching the glass/solution interface is deflected into the solution. The intensity of the incident light that remains in the glass is dependent on the difference between the refractive indexes of the glass and the solution. The method is insensitive to the presence of the crystals up to a slurry density of about 30% crystal mass. However, the measurement is sensitive to ambient light and the presence of gas bubbles that may adhere to the surface of the refractive reference material (the glass tube in the previous case). It is interesting to note that Nývlt and coworkers measure similar but not identical supersaturation curves using conductive and refractive measurements. They comment that the scatter in the conductivity based measurements is "significantly larger" than for the refractive method. However, this reduced scatter is not clear from their comparative plots. Mullin and Leci [98] used refractometry to investigate the crystallization of citric acid from water. They report a measurement accuracy of $\pm 2 \times 10^{-3}$ grams of citric acid monohydrate per gram of water with 30 to 60 second sampling times.

A final concentration measurement technique, attenuated total reflection Fourier transform infrared spectroscopy (ATR FTIR), has been developed by Berglund and coworkers [35, 36, 24]. This method works by measuring changes in the infrared spectrum of a solution. First, the peaks in an infrared spectrum of the pure solvent are identified. This spectrum is compared to a similar spectrum of the solute-solvent system. The new spectrum contains the same solvent peaks plus some new peaks that correspond to the reflection spectrum of the dissolved solute molecules. If the solute peaks are at different wavelengths than those of the solvent, the solute concentration may be inferred from a ratio of the height of a solute peak to the height of a solvent peak. The sensitivity of the technique depends upon the sensitivity of the peak heights to the solute concentration. The method may be used *in situ* because the penetration depth of the infrared signal is small relative to the fluid boundary layer around the sensor. Penetration depth is a function of wavelength, angle of incidence, and the relative difference between the refractive indexes of the solution and the signal propagating medium. This method may also be used to monitor concentration changes for more than one solute and may therefore be used in tertiary systems. Densitometry is restricted to binary systems. The method is ineffective if there are no distinct solute peaks. Dunuwila and Berglund [35] demonstrate the use of the technique to determine solubility curves and track supersaturation profiles for the aqueous crystallization of Maleic Acid.

Though the precision of densitometry justifies its use, the measurement is troublesome for cases in which phase separation is complicated by incrustation. Of the technologies available, ATR FTIR appears to provide the best combination of convenience, precision, and flexibility. ATR FTIR also offers the advantage of allowing measurements in multiple component systems, which may be useful for detailed studies of crystallization kinetics in systems with high levels of impurity or in systems where driving force is generated by the combined effects of temperature and non-solvent addition. Unfortunately, ATR FTIR systems are an order of magnitude more expensive than competing techniques like densitometry.

2.5.3 Transmittance Measurement

This section describes the transmittance measurement that provides this study's measurement of the second moment of the CSD. Due to theoretical restrictions, this measurement may only be accurately related to the second moment when slurry densities are very low. Thus, this measurement is used exclusively in the early stages of crystallization runs and is the primary measure of the early time solid phase dynamics in the system. The measurement is performed by an *in situ* colorimeter probe.

Theory Development

Transmittance is defined as the fraction of incident light that may be transmitted through a slurry sample:

$$\text{transmittance} = \frac{I}{I_0} \tag{2.2}$$

where I is the intensity of the light transmitted through the slurry and I_0 is the intensity of the incident light. This type of measurement may be performed by many types of instruments including light scattering particle sizers, spectrophotometers, and colorimeters. This study employs a visible light colorimeter (Brinkmann Instruments, model PC-801) on loan to the University of Wisconsin–Madison from the Eastman Kodak Company.

The transmittance measurement is related to the slurry turbidity by the Beer-Lambert Law:

$$\frac{I}{I_0} = \exp(-\tau_s l) \tag{2.3}$$

where τ_s is the slurry turbidity and l is the beam length of the slurry sample through which the light is transmitted. The turbidity may be further related to the CSD:

$$\tau_s = \int_0^\infty Q(L) f(L) A_p(L) dL \tag{2.4}$$

where L is the characteristic length of a crystal, $f(L)$ is the CSD, $Q(L)$ is the extinction efficiency factor, and $A_p(L)$ is the projected area for a crystal of characteristic size L. The expression may be simplified given further knowledge of the forms of $Q(L)$ and $A_p(L)$.

The extinction efficiency factor describes the reduction of the intensity of the incident light due to the interference of a particle located in its path. Extinction is a function of the type of light scattering that is occurring in the system. For most crystallization applications, including this one, the crystals studied are large relative to the wavelength of the incident light. The monitored wavelength of the incident light used in this study is 0.670 microns. For this “large particle” case, the dominant light scattering mechanism in the “near-forward” direction is Fraunhoffer diffraction. Assuming that the scattering is due solely to Fraunhoffer diffraction, the scattering is independent of refractive index, depends only on the projected area of the particles, and has an extinction efficiency factor of 2 for all particle sizes. The energetic implication of $Q(L) = 2$ is that a particle removes from the incident beam twice the amount of light that it blocks due to its projected area. The effective doubling of this area is the result of the destructive interaction of light diffracted by the particle with incident light. Discussions of extinction for both the large and general particle cases are given by van de Hulst [148].

To account for the projected area of non-spherical crystals, Brown and Felton [18] suggest corrections using a theorem proved by van de Hulst [148]. The theorem states that the average geometrical cross section of a randomly oriented convex particle is equal to one-fourth its total surface area. Given that the surface area of a crystal may be written as proportional to the square of its characteristic dimension:

$$A_s = k_a L^2 \tag{2.5}$$

where k_a is the area shape factor, the theorem result may be combined with Equation 2.5 and an extinction efficiency factor of 2 to rewrite Equation 2.4:

$$\tau_s = \int_0^\infty 2f(L)\frac{k_a L^2}{4} dL = \frac{k_a}{2}\int_0^\infty f(L)L^2 dL \tag{2.6}$$

where $\int_0^\infty f(L)L^2 dL$ is the second moment of the CSD per unit volume, often written as μ_2. Recalling Equation 2.3, the log of the transmittance measurement is proportional to the second moment of the CSD. k_a is determined from scanning electron micrographs of the crystals. Appendix A provides a description of the method used to determine k_a and k_v, the volume shape factor, for the seed crystals.

A primary assumption in the preceding development is that the measurements are limited to a single-scattering operating regime. Single scattering implies that disturbances due to additional particle interference with previously scattered light are negligible. In practice, the robustness of the transmittance measurement allows some extension out of the regime of single scattering. Erroneous measurements occur when light scattered out of the path of the incident light is scattered back into that path by other particles. The errors may be positive due to the added intensity of the multiply-scattered light [105] or negative due to phase-related destructive interference of that light with incident light [95]. Intuitively, the probability of significant scattering of light back into the exact path of the incident light seems low if there is only one additional scattering event (double scattering). Thus, in practice, transmittance measurements are extended into regimes of double scattering. This restriction still limits the use of the transmittance measurement to dilute slurry samples and has motivated the development of on-line dilution schemes by some researchers [74].

Van de Hulst [148] presents heuristic rules describing scattering regimes as a function of transmittance: single scattering is prevalent for transmittance values higher than 0.90; with double scattering from 0.90 to about 0.75; and multiple scattering below 0.75. Boxman [17] recommends an operating range of $0.75 < \frac{I}{I_0} < 0.90$, identical to van de Hulst's double scattering regime, to maximize signal-to-noise ratio for transmittance measurements. Miller [95] and Witkowski [158] each used transmittance values greater than about 0.70. In this study, due to smaller crystal sizes and the needle-like habit, a more conservative minimum value of 0.80 was used.

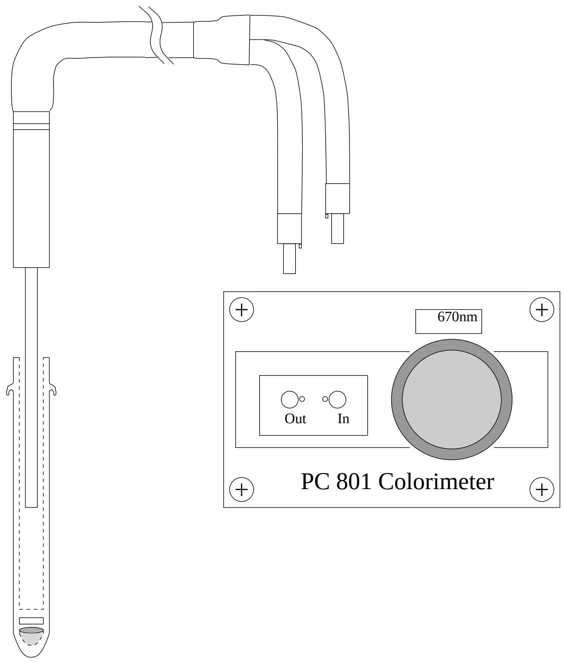

Figure 2.4: Transmittance probe tip with fiber-optic probe and colorimeter light filter.

Colorimeter Apparatus

Unlike the densitometry based concentration measurement, an external sample line is not required for the transmittance measurement. The colorimeter uses an *in situ* probe that is seated in an 18 millimeter port in the headplate of the crystallizer. The probe, illustrated with its sensor and Pyrex tip in Figure 2.4, is a 4 millimeter inside diameter stainless steel tube containing two bundles of optical fibers. The 30 inch optical fibers run from the end of the probe to the colorimeter. One bundle transmits visible light from the colorimeter light source to the end of the probe while the other receives light at the probe and transmits it back to the colorimeter's light filter and sensor. As noted earlier, the filter is set to allow the source to monitor light with a wavelength of 670 nanometers.

A Pyrex glass tip fits over the end of the probe and is held in position by a Vyton fastener. The tip encloses the probe to protect it from direct contact with the process fluid.

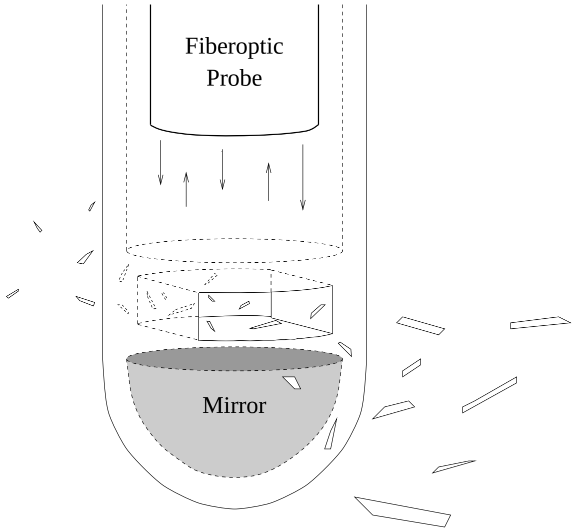

Figure 2.5: Close-up view of colorimeter probe and tip.

At the end of the tip there is a glass window, a slot through which the slurry may flow, and a mirror to reflect the light from the source fibers back along its incident path and into the receiving fibers. This arrangement is shown in Figure 2.5. The slot allows a constant sample of slurry to flow across the path of the light. The slot is 1 millimeter in depth and 4 millimeters in width. The effective beam length is 2 millimeters as the light path crosses the depth of the slot twice, once directly and once after reflection in the mirror.

The colorimeter has analog outputs to allow the transmittance value to be collected directly by the data acquisition hardware. However, because crystals and air bubbles are occasionally trapped in the probe slot, the data are read as percent transmittance from the colorimeter's digital display and written into the experiment notebook by hand. This allows the experimenter to check that the slot is clear before taking a measurement. When necessary, the slot is easily cleared by gently jostling the probe. Data points are collected at intervals of approximately two minutes and dead time of the measurement is negligible.

Transmittance Validation

The transmittance measurement of the colorimeter may be validated experimentally. For a particulate system of known particle shape and size distribution, the Beer-Lambert law may be used to determine the expected value of the transmittance given the mass of the particulate phase and the volume of the continuous phase. Comparison to transmittance data collected for the system then provides a measure of the adequacy of the device in question. The NaCl–water system provides an adequate model system.

NaCl crystals have a size independent, cubic shape and may be separated into distinct size classes by sieving. The system is particularly easy to hold at the saturation point (i.e. no crystal growth or dissolution) because the temperature dependence of the solubility curve is very weak over a wide range. A temperature change from 20.0°C to 40.0°C results in only a 1.7% increase in solubility [107].

A saturated solution was prepared by massing appropriate amounts of NaCl and deionized water and then heating the slurry until all of the crystals were dissolved. Make-up water was added to the system to account for evaporation during the heating process. A measured volume of this saturated solution served as the continuous phase for the validation experiment. The solid phase sample consisted of sieved NaCl crystals. These crystals were between 250 and 300 microns in size. By making the assumption that these crystals had a size of 275 microns, the number of crystals in a massed sample may be determined by dividing the mass of the sample by the mass of a single crystal:

$$N = \frac{W}{\rho_c L^3} \tag{2.7}$$

where W is the mass of crystals and ρ_c is the crystal density. Equation 2.7 may be com-

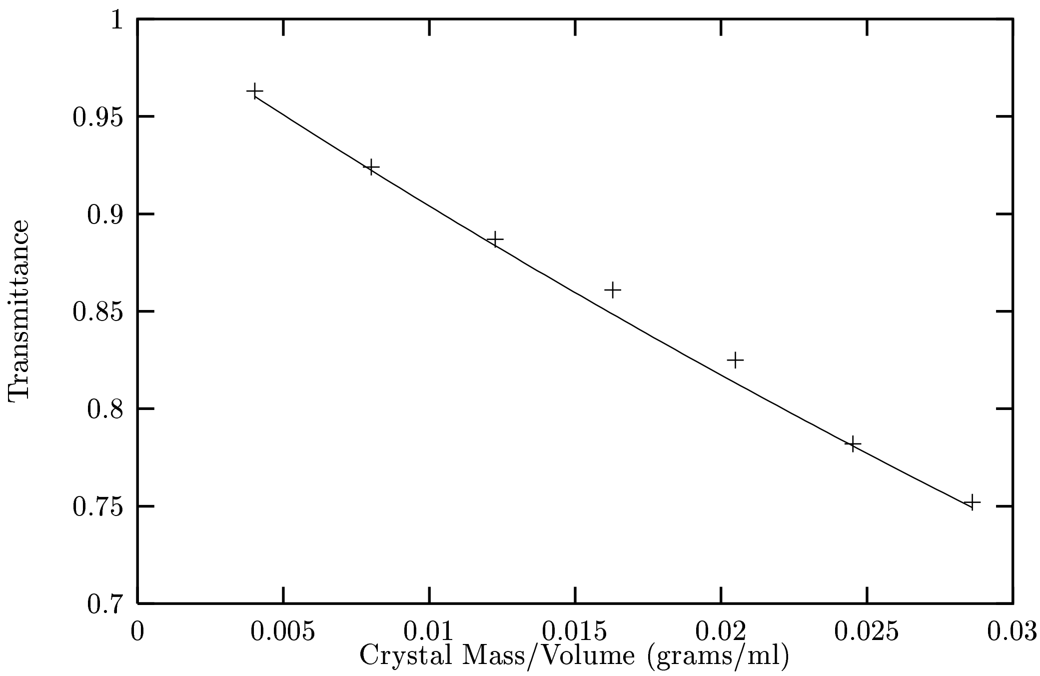

Figure 2.6: Transmittance validation data collected for a mono-dispersion of 275 micron, cubic NaCl crystals in a saturated aqueous solution.

bined with Equations 2.3 and 2.6 to provide an analytical expression that predicts the transmittance for a mono-dispersion of cubic crystals in a known volume, V, of continuous phase.

$$\frac{I}{I_0} = \exp\left(-\frac{3Wl}{\rho_c LV}\right) \tag{2.8}$$

Note that the area shape factor, k_a in Equation 2.6, is equal to 6 for cubic crystals and that l is the beam length of the instrument.

Figure 2.6 compares colorimeter transmittance measurements to Equation 2.8 for known masses of NaCl per unit volume of saturated solution. The plot indicates that for an idealized system the Beer-Lambert law provides a good description of the colorimeter transmittance data over the range of interest.

2.6 Filtration Measurements

A primary goal of this study is to enhance the performance of the solid-liquid separation processes that follow batch crystallization on a production line. In order to achieve this goal a metric must be defined that allows enhancements to be quantified and thereby allow the comparison of different operating policies. Such a measure is especially important to this study as the small size and irregular crystal habit prevent experimental verification of the actual shape of the CSD. The chosen metric, defined in the following section, is the average specific cake resistance, $\overline{\alpha}_c$. The cake resistance is chosen over the more intuitive measure of filtration time because it is not a function of the solid phase density or the total mass of the filtered slurry. Thus, it allows comparison of a wider variety of product slurries and removes some disturbances concerning the replicability of the preparation of the filtration slurries. An additional factor in the choice of average specific cake resistance is its use by the Chemical Engineering Research Laboratory (CERL) at the Eastman Kodak Company. Thus, cake resistance provides a measure of comparison between the ongoing academic and industrial studies of this chemical system.

2.6.1 Definition of Specific Average Cake Resistance

The Ruth modification of Darcy's Law for one-dimensional flow in a porous media may be written as

$$\frac{dP_l}{dw} = \alpha_c \mu_f q_f \tag{2.9}$$

where P_l is the hydraulic pressure across the height of the filter cake, w is the mass of crystal deposited on the filter medium per unit area, α_c is the specific cake resistance, μ_f is the viscosity of the filtrate, and q_f is the flow-rate of liquid phase through the filter cake per unit area. Integrating Equation 2.9 over pressure and mass of crystal, the average specific cake resistance is defined as

$$\frac{1}{\bar{\alpha}_c} = \frac{1}{P - P_m} \int_0^{P-P_m} \frac{dP_l}{\alpha_c} = \frac{\mu_f q_f w}{P - P_m} \tag{2.10}$$

where P is a constant pressure applied at the top of the cake and P_m is the pressure required to overcome the resistance of the filter medium. By redefining some terms:

$$\begin{aligned} \nu_f &= \frac{\mu_f}{\rho_f} \\ q_f &= \frac{1}{A_f} \frac{dm_f}{dt} \\ P_m &= \frac{R_m \nu_f}{A_f} \frac{dm_f}{dt} \end{aligned}$$

where ν_f is the kinematic viscosity of the filtrate, m_f is the mass of the filtrate (filtered liquid phase assumed to contain no crystals), A_f is the cross sectional area of the filtration apparatus, and R_m is the flow resistance of the filter medium, Equation 2.10 may be recast:

$$\frac{dt}{dm_f} = \frac{\bar{\alpha}_c \nu_f}{A_f P} w + \frac{R_m \nu_f}{A_f P} \tag{2.11}$$

Recalling that w is defined as the mass of cake deposited per unit area, we may make a substitution for w:

$$w = \frac{c_f}{A_f} m_f \tag{2.12}$$

where c_f is defined as the ratio of the mass of solid phase to the mass of continuous phase in the filtered slurry. This substitution assumes that the mass of the cake is proportional to the mass of filtrate that has passed through the cake. Substituting Equation 2.12 into Equation 2.11 and integrating over time and mass of filtrate one finds a linear relationship

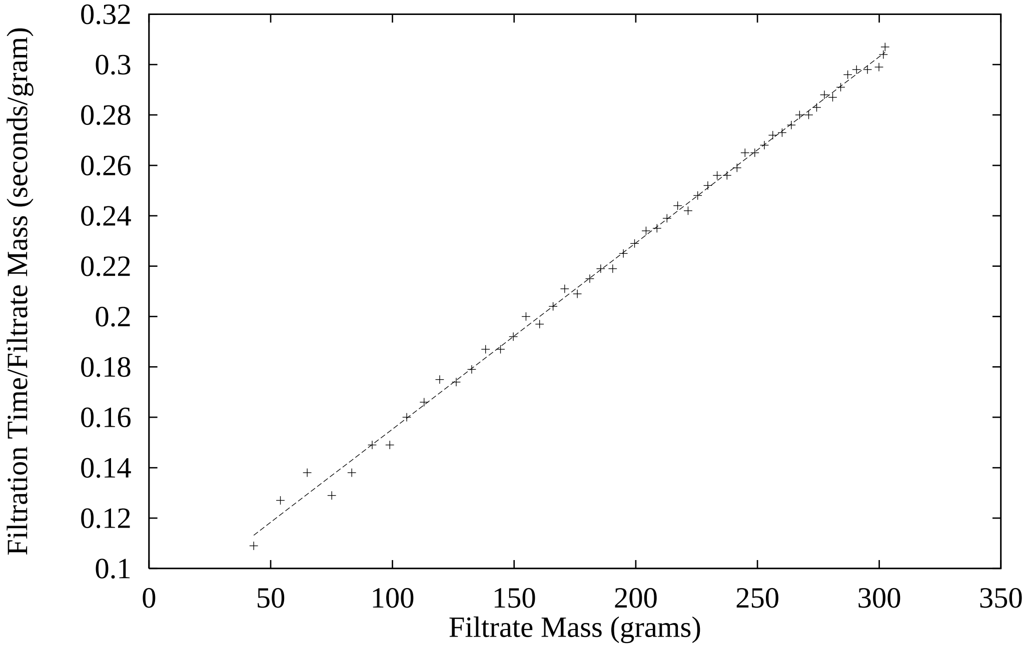

Figure 2.7: Plot of t/m_f versus m_f for a typical constant pressure filtration run. $\bar{\alpha}_c$ is determined to be 2.49×10^9 cm/g in this example.

between t/m_f and m_f:

$$\frac{t}{m_f} = \frac{\bar{\alpha}_{\mathbf{c}} \nu_f c_f}{2A_f^2 P} m_f + \frac{R_m \nu_f}{A_f P}. \tag{2.13}$$

By collecting m_f versus time data for a particular filtration, the average specific cake resistance may be determined from the regressed slope of t/m_f versus m_f data and the resistance of the filter medium may be determined from the regressed intercept. An example plot of a linear fit to this type of data is given in Figure 2.7. The preceding development follows that of Tiller and Crump [145] for the case of a constant pressure filtration experiment.

2.6.2 Filtration Apparatus

The filtration apparatus used in this study is a constant pressure unit produced by Millipore Corporation. An exploded view of the unit is shown in Figure 2.8. The inside diameter of the barrel is 7.7 centimeters, corresponding to a filtration area of approximately 47 square centimeters. Two components not included in the illustration are a teflon o-ring that seals the contact between the headplate and the barrel and the filter medium itself. This study uses a Dacron weave filter medium that is cut to size and installed on top of the filter screen shown in Figure 2.8. The Dacron filter material was provided by Eastman Kodak Company and is used in some of the Company's production filtrations. The apparatus is held together by three posts that connect the base and headplates. Pressure applied by the nuts that secure these posts compresses the o-rings and seals the system. With the exception of the o-rings, the filter medium, and the polycarbonate barrel, all components are made of stainless steel.

In typical operation, approximately 500 milliliters of slurry is pumped from the crystallizer into the filter apparatus by a peristaltic pump (Masterflex, model 7520-25). The pressure valve is then closed and a constant pressure source is attached to the port in the headplate. This source is provided by a compressed air line, constrained to a pressure of approximately 5.5 psig by a regulator (Speedaire, model 2Z767D). In practice, pressures varied from 4.6 to 7.1 psig for different experiments but were constant within each experiment. When constant pressure is applied the outlet valve is opened, starting the filtration. Initially, the filter medium alone filters the crystals from the slurry. However, a cake of filtered crystals forms quickly on the medium and becomes the *de facto* filter. It is at this point that the equations developed in Section 2.6.1 become applicable. The solids-free filtrate flows through the cake to the outlet valve and into an Erlenmeyer flask positioned on an electronic balance (Mettler, model PM3000). The balance records the filtrate mass at intervals of approximately 2 seconds, sending the readings to the serial port of a personal computer (IBM, model 7531, 80286 cpu). There the signal is read, given a time stamp, and written to file by a data acquisition program (MASSLOG, version 2.10, written by Todd Colin of the Eastman Kodak Company). This data is subsequently used to generate plots similar to that in Figure 2.7. A step-by-step description of a constant pressure filtration test is given with the experimental procedures of Section 2.8. Details concerning the balance settings required for communication with the MASSLOG program are given in Appendix A.

2.7 Data Acquisition

Data from the various measurement devices described in previous sections are automatically collected by a data acquisition system. The primary components of the system are the VAXStation computer that performs on-line calculations and stores the data and the signal conditioning system. In addition to collecting, monitoring, and archiving the data, the data acquisition system can pass control set points back to particular instruments in the crystallizer apparatus. This functionality allows the system to control the crystallizer and trace bath temperatures. This section provides a detailed description of the hardware and software elements of the data acquisition system.

2.7.1 VAXStation Communications

A Digital Equipment Company (DEC) VAXStation II using the VMS 5.5-2 operating system is the primary link between the operator and the crystallization system. In addition to collecting, storing, and displaying the on-line data, the computer runs the model predictive controller described in Section 2.4.2. The VAXStation has been modified to perform these tasks by the addition of three DEC communication boards; the AXV11-C analog input/output board, the AAV11-C analog output board, and the DZQ11 asynchronous multiplexer board.

The AXV11-C board allows the VAX to receive temperature data and communicate with the Omega PID controller and the Haake bath (see Section 2.4.1). The board converts 16 single-ended, 0-10V input signals to 12-bit digital signals and has two digital-to-analog converters capable of converting 12-bit digital outputs to 0-10V analog signals. The AAV11-C provides the VAX with two additional digital-to-analog converters, giving the system the capacity to control a total of four analog devices. At present, the AAV11-C is not in use as the system only requires two control outputs. DEC software device drivers in the Data Translation subroutine library of the MicroVMSLIB assist in control and communication with the AXV11-C and AAV11-C boards. The FORTRAN routines that store the data and perform the control computations communicate with the boards using subroutines included in this library. The original MicroVMS drivers were modified by previous researchers [158, 95] to allow them to operate under the full VMS operating system.

The DZQ11 board lets the VAX communicate with devices that use RS-232C interfaces. The DZQ11 has four ports that may be configured individually to conform with the transmission rates and formats of different devices. The ports are configured using the VMS "set term" command. FORTRAN routines can communicate with the board and the measurement devices using VMS utility routines for port-to-channel assignments (SYS$ASSIGN) and queued input/output (e.g., SYS$QIOW). Both the densitometer and the Omega PID controller communicate with the VAX via RS-232C interfaces.

2.7.2 Signal Conditioning System

Particular measurement devices may send or receive analog signals that are inappropriate for direct communication with the VAXStation boards. For this reason, a signal conditioning system is used to convert such signals to the proper forms. The 3B Series Signal Conditioning I/O Subsystem produced by Analog Devices Incorporated (ADI) performs this function for the crystallizer data acquisition system. The system consists of a series of modules that plug into a mounted backplane. Each module corresponds to a particular analog device and performs the signal conversion required to allow that device to communicate with the VAXStation. RTD modules (model 3B34) generate excitation signals for the RTD's and then convert the return signals to 0-10V signals that can be received by the VAXStation's AXV11-C board. These same modules send 4-20 milliamp signals to a light emitting display (LED) that allows the operator to visually monitor the temperature data. A similar module (model 3B32) converts output signals from the Omega PID controller to the 4-20 milliamp signal required by the 3-way control valve.

2.8 Operating Procedures

This section presents step by step operating procedures for the crystallizer and the filtration apparatus.

2.8.1 Crystallizer Operation

- The crystallizer is charged with initial loads of solvent and solute constituting the desired solute concentration. Typical charges were 154 grams of photochemical and

1400 grams of heptane. These were measured on a high capacity electronic balance (Mettler, model PM16).

- The charged crystallizer is placed in the fume hood and connected to the jacket water and densitometer lines. The crystallizer temperature RTD and the stir motor are attached and the crystallizer is wrapped in foam rubber insulation.

- The jacket water valve is opened and the gear pump, PID controller, and 3-way control valve are switched on. In order to dissolve the solute, the jacket controller set point is approximately 5°C above the saturation temperature of the solution. For the typical charge, this corresponds to a dissolution temperature of approximately 40°C.[3] The Advantage water bath (heat source for the jacket water) is set to 60°C to achieve this jacket temperature. The crystallizer is maintained at this temperature for approximately 15 minutes after dissolution appears to be complete.

- During the 15 minute post-dissolution period, the colorimeter probe is installed in the headplate of the crystallizer and the densitometer pump is turned on. The Haake baths are operated at set points of 35°C. The colorimeter is turned on to allow the instrument to warm up before the start of the run.

- The crystallizer is cooled to the desired initial temperature by manually adjusting the PID controller. The Advantage heater set point is reduced to 35°C and the Haake baths are adjusted to the appropriate values (see Section 2.5.2). If necessary, ice is added to the cold-side jacket baths to provide sufficient cooling capacity. Once

[3] The photochemical material used to charge the crystallizer is a different polymorph than the form crystallized in this apparatus. The saturation temperature for the initial form is approximately 10°C higher than for the crystallized form.

the temperature of the crystallizer is within 0.5°C of the desired initial temperature, the MPC controller is switched on to move the process to the exact set point. This temperature is maintained for approximately 15 minutes.

- During the holding period, the probe tip and densitometer lines are cleared of any air bubbles. The colorimeter reading is adjusted to read 100% transmittance using fine and coarse adjustment knobs. The trace bath for the densitometer return is switched from internal to external set point tracking. A file defining the desired crystallizer temperature trajectory is prepared for the MPC control program.

- The seeds for the crystallization are sampled from the filtered product of a previous run. The mass of the seed load is measured on a electronic balance with a 0.0001 gram precision (Mettler, model AE 163). Seed loads in this study varied from 0.05-0.25 grams.

- The start command is given to the VAXStation, initiating the MPC temperature controller and the data collection programs. The initial temperature is maintained for one minute, at which time the seeds are added.

- If incrustation of the densitometer filter is not a concern, the process requires only that the operator record transmittance data (see Section 2.5.3). Otherwise, densitometer measurements may be collected intermittently or a back-flush strategy may be employed (see Section 2.5.2). Such densitometer schemes require the operator to make manual adjustments of the densitometer pump speed and direction. Once the measured transmittance is below its minimum value, the colorimeter probe may be removed at any time.

- Once a run is completed, the stop command is given to the VAXStation to halt controller and data acquisition programs. The gear pump, PID controller, and 3-way control valve are switched off and the jacket water valve is closed. The outlet stream to the densitometer is disconnected but the pump continues running until all the liquor in the line is returned to the crystallizer. The stir motor and RTD are removed and the densitometer return and jacket water lines are disconnected. Rubber plugs replace the jacket water lines, keeping the jacket full to insulate the crystallizer until to slurry filtration can be performed.

- The full insulated crystallizer is moved to another fume hood for filtration while the densitometer is cleaned by recirculating clean hexane through the sample loop. After filtration, the crystallizer, headplate, and ports are cleaned with acetone. The densitometer filter can be disassembled to facilitate cleaning or to changed the wire mesh.

2.8.2 Filter Operation

- Prior to filtration, the filter medium is cut to size and installed in the filtration apparatus. The sealed filter apparatus (outlet and pressure valves closed) is pressure tested with compressed air to check for leaks. The pressure release valve is then opened to allow slurry to be pumped through the port in the headplate. The MASSLOG program is initiated on the PC. This involves the specification of a data file and graphical display options. A clean Erlenmeyer flask is placed on the balance and positioned to receive the filtrate. The balance is zeroed.

- If the crystallization slurry has a sufficient solids density the slurry can be pumped directly into the filter apparatus.[4] Otherwise, the slurry is allowed to settle and clear liquor is removed until the density is sufficient. Typical settling times were 8 minutes followed by the removal of approximately 500 grams of clear liquor.

- The slurry is pumped into the apparatus through a 1/4 inch stainless steel tube that is simultaneously used to stir the slurry. The mass of a typical slurry sample is approximately 340 grams, with a corresponding volume of about 500 milliliters.

- The regulated compressed air line is connected to the port in the filter headplate and the pressure release valve is closed. The air flow in the fume hood is cutoff to minimize noise in the filtrate mass measurement. The MASSLOG program is started.

- The valve for the compressed air line is opened and, after a short pause, the filter outlet valve is opened and filtration begins.

- The operator records the pressure applied to the slurry during the filtration. This pressure may drift slightly but should be approximately constant. The filtration experiments were designed to operate at a pressure of 5.5 psig. Once breakthrough is achieved, the pressure release valve is opened and the compressed air valve is closed.[5] The MASSLOG program is turned off having recorded the final filtrate mass.

- The filter apparatus is disassembled and the wet cake is massed and set in the reactivated fume hood to dry. All components of the apparatus are cleaned with hexane

[4] A mass density $> 5\%$ solids is preferred to maximize the amount of data to be collected in the linear region depicted in Figure 2.7. Smaller mass densities may result in shorter filtration times and noisier data.

[5] Breakthrough denotes the point at which the liquid-air interface reaches and passes through the filter cake.

(acetone may damage the polycarbonate barrel).

- After drying overnight, the filter cake is remassed to obtain the crystal mass in the filtered slurry sample. Some chemical systems may require a longer drying time.

2.9 Summary

This chapter has presented a detailed review of the batch crystallization apparatus used in this study. The main apparatus consists of a jacketed glass batch cooling crystallizer equipped with an array of measurement devices; an external densitometer for on-line concentration measurement, an *in situ* colorimeter probe for on-line transmittance measurement, and a set of RTD's to monitor various temperatures in the system. The apparatus also includes a temperature control system for the jacket cooling water. This system consists of heating, cooling, and pump hardware and a model-based controller that runs on a VAXStation computer. Control signals and data collection are performed by a data acquisition system that includes both the VAXStation and a signal conditioning board.

A constant pressure filtration apparatus has also been described. This apparatus consists of a sealed slurry reservoir (500 milliliter capacity) and a filter medium support structure. Filtration pressure is provided by a regulated line of compressed air. The mass of filtered liquor is measured by an electronic balance that communicates via serial port with an IBM PC that records on-line mass versus time data. Data from this apparatus is used to determine the average specific cake resistance of the crystallized slurries.

The crystallizer and filter have been designed to provide experimental data that are sufficient to identify and validate the crystallization models developed later in this thesis. The filtration apparatus is of particular use in confirming the success or failure of this study's efforts to optimize crystallizer operating procedures to facilitate the filtration of final-time slurries. The photochemical-heptane system used in this study is a relevant industrial system currently produced by the Eastman Kodak Company.

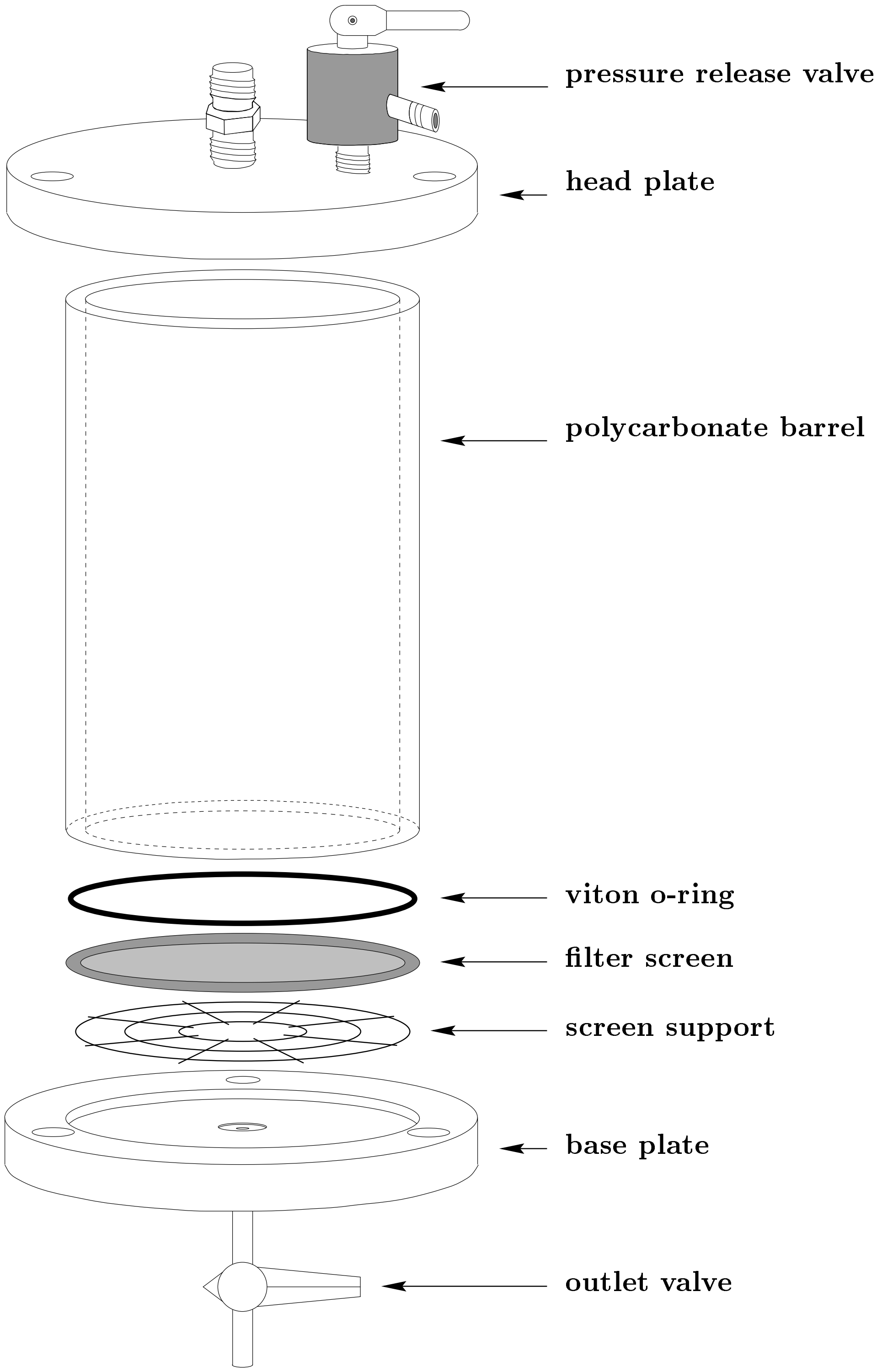

Figure 2.8: Exploded view of constant pressure filter apparatus.

Chapter 3

Model Formulation and Solution

The process of model identification for a particular physical system is comprised of two interrelated parts. First, a general model structure must be defined. This structure needs to account for the pertinent states of the system (e.g. concentration, temperature, CSD) and include any additional functions that may drive these states (e.g. crystal nucleation and growth). The general structure must also employ response models to map the states into the space of the data. The response models allow for the prediction of the measured behavior of the system and, in so doing, are the keys to the second phase of model identification, the estimation of model parameters from experimental data.

The two steps of identification comprise an iterative process. A model is proposed, data is collected and parameters are estimated. If the there are no admissible parameter sets that allow the model to describe the data, the model may be revised and the process repeated.[1] This chapter describes the development of a model for the batch crystallizer apparatus described in Chapter 2. The numerical solution methods required to solve the

[1]This presumes that the data are infallible. Since this is seldom the case in practice, the critical examination of the data through replication is also necessary. Thus, model identification is actually a three-part process involving the model, the experimental apparatus, and the estimation of optimal parameters.

model are also presented. The estimation of the kinetic parameters in the model is reserved for Chapter 4.

Chapter Outline. The batch crystallizer model uses a population balance equation (PBE) to describe the evolution of the crystal size distribution during a crystallization experiment. The general form of the PBE is derived in Section 3.1.1 and is simplified to describe seeded batch crystallizers in Section 3.1.2. The PBE is augmented by mass and energy balances that model the dynamics of the solute concentration in the solution phase and the system temperature, respectively. These models are presented in Section 3.2 with the full derivations provided in Appendix D. Batch crystallization is driven by the kinetic processes of crystal nucleation and growth. Empirical power law models of these processes, along with a simple dissolution model, are given in Section 3.3. The response models for the general model are the transmittance model of Section 2.5.3 and the solution phase concentration provided by the mass balance. The batch crystallizer model is summarized in Section 3.4.

The PBE is a partial differential equation in time and crystal size. Therefore, the crystallizer model must be solved numerically. General solution methods for this type of model are described in Section 3.5. This study solves the PBE via the method of moments, which allows the reduction of the PBE to a set of four ordinary differential equations (ODE). The moment method is presented in Section 3.6. The implementation of the model and the specification of the initial conditions for the moment equations are presented in Section 3.7. A brief simulation study introducing some interesting aspects of the model is also presented in Section 3.7.

3.1 Modeling of Particulate Processes

The motivation for modeling a process is to gain an accurate mathematical representation of the behavior of the physical system. Given such a representation, tasks such as simulation, optimization, and control may be explored. Dispersed phase systems, like crystallization, are particularly challenging to model because the behavior of the system depends not only on the environmental variables (concentration, temperature) but is also a function of certain qualities of the individual particles. These qualities are called the characteristic variables. The characteristic variables that influence performance vary from system to system. For example, in a cell fermentation process the developmental stage of a cell may play a role in the cell's production rate of an antibiotic or other desired product. Optimizing production in such a process requires a knowledge of the distribution of the ages of the individual cells in the system [114, 45]. Other characteristic variables in particulate systems include: micelle size in emulsion polymerization [125, 126], catalyst particle activity or particle size in multiphase reactors [132, 7, 30], and crystal size or growth rate in solution crystallization [123, 57]. A mathematical framework suited to modeling these types of particulate systems is the population balance.

3.1.1 Deterministic Population Balance

The population balance concept was first proposed in 1962 by Rudd [132] as a generalization of residence time distribution to a characteristic variable other than particle age. In his case the describing variable was catalyst particle activity in continuous catalytic reactor-regenerator systems. Rudd's work was followed by several researchers [120, 30, 7] who explored the application of the method in crystallization and reactive mixing processes. The most influential of the these works was published by Hulbert and Katz [71]. These authors extended the population balance to include an arbitrary number of characteristic variables and developed solution methods for a variety of potential applications using the method of moments.

The deterministic population balance equation (PBE) may be derived as a conservation equation for the number of particles in a population. Given a population of particles suspended in a continuous fluid, let $f(\boldsymbol{z},t)$ be the number of particles per unit volume of the phase space, $\boldsymbol{z}$. $f(\boldsymbol{z},t)$ is a time-dependent particle density analogous to the mass density, $\rho(\boldsymbol{x},t)$, of a compressible fluid. However, whereas $\rho(\boldsymbol{x},t)$ defines the distribution of fluid density over time and the conventional spatial coordinates, $\boldsymbol{x}$, $f(\boldsymbol{z},t)$ defines the particle density over time, the spatial coordinates, and the set of characteristic coordinates, $\boldsymbol{y}$.

$$\boldsymbol{z} = \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix}$$

The $\boldsymbol{y}$ and $\boldsymbol{x}$ coordinate sets are termed the internal and external coordinates, respectively. This emphasizes that while the external coordinates define the spatial location of a particle, the internal variables are qualities, such as size or purity, that distinguish individuals by virtue of differences in internal structure. Often, as is the case in crystallization, the macroscopic behavior of a particulate system depends on how the particles are distributed over the space of the internal variables.

Let Ω be an arbitrary subregion of the combined phase space of the internal and external coordinates and let $d\Omega$ be an infinitesimal volume within that subregion. The total

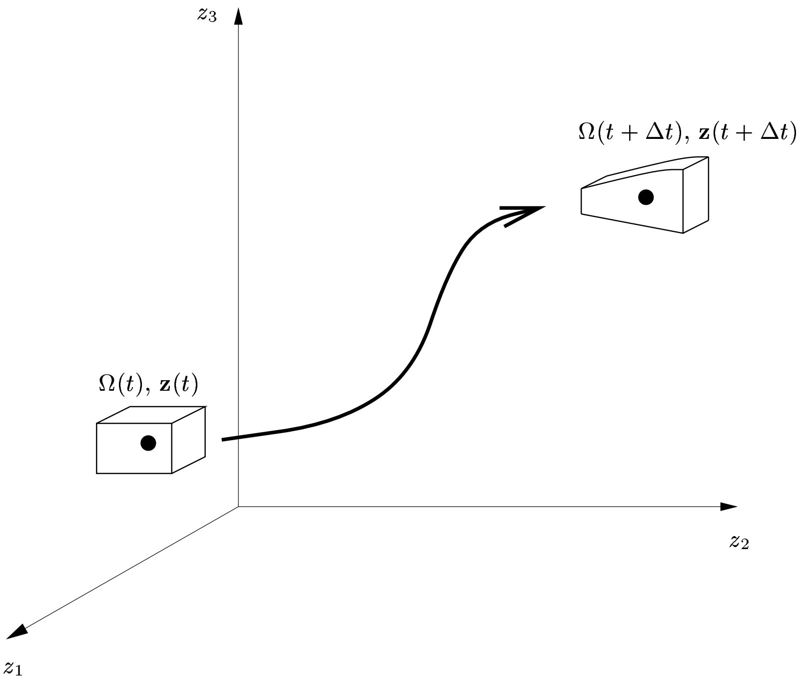

Figure 3.1: Phase space movement and dilation of an arbitrary volume element Ω over time.

number of particles within subregion Ω is equal to

$$\int_{\Omega} f(\boldsymbol{z},t)d\Omega. \tag{3.1}$$

The material derivative of this particle count depends on the rates of birth, B, and death, D, for particles in Ω.

$$\frac{d}{dt}\int_{\Omega} f(\boldsymbol{z},t)d\Omega = \int_{\Omega}(B-D)\,d\Omega \tag{3.2}$$

Assuming a non-equilibrium environment, the population of particles inside Ω moves through the phase space over time. As it does so, the surface of Ω expands or contracts to contain the population. This process, termed dilation, is depicted in Figure 3.1 for 3-dimensional $\boldsymbol{z}$. It is described mathematically by rewriting the left side of Equation 3.2 as

$$\frac{d}{dt}\int_{\Omega} f(\boldsymbol{z},t)d\Omega = \int_{\Omega}\frac{\partial f(\boldsymbol{z},t)}{\partial t}d\Omega + \int_{\Sigma} f(\boldsymbol{z},t)\,(\boldsymbol{v}\cdot\boldsymbol{n})\,d\Sigma \tag{3.3}$$

where Σ is the surface of the subregion Ω, $\boldsymbol{v}$ is the vector of particle velocities corresponding to the coordinates of the phase space, and $\boldsymbol{n}$ is the outward normal perpendicular to Σ. The divergence theorem is used to convert the surface integral term to a volume integral.

$$\int_{\Sigma} f(\boldsymbol{z},t)\,(\boldsymbol{v}\cdot\boldsymbol{n})\,d\Sigma = \int_{\Omega} \nabla\cdot(f(\boldsymbol{z},t)\boldsymbol{v})\,d\Omega \tag{3.4}$$

Substituting Equations 3.3 and 3.4 into 3.2 and consolidating all terms in a single integration provides

$$\int_{\Omega}\left[\frac{\partial f(\boldsymbol{z},t)}{\partial t} + \nabla\cdot(f(\boldsymbol{z},t)\boldsymbol{v}) - (B-D)\right] d\Omega = 0. \tag{3.5}$$

Since Ω is arbitrary, the integral vanishes identically providing the general microscopic population balance.

$$\frac{\partial f(\boldsymbol{z},t)}{\partial t} + \nabla\cdot(f(\boldsymbol{z},t)\boldsymbol{v}) = B - D \tag{3.6}$$

The microscopic population balance may be further simplified for ideally-mixed systems in which the particles have no spatial dependence. Recall that the vector $\boldsymbol{z}$ is partitioned into the internal and external coordinates, $\boldsymbol{y}$ and $\boldsymbol{x}$ respectively. Since f, B, and D have no spatial dependence for the well-mixed case, Equation 3.6 may be integrated over $\boldsymbol{x}$ to give the following,

$$V\frac{\partial f(\boldsymbol{y},t)}{\partial t} + V\nabla\cdot(f(\boldsymbol{y},t)\boldsymbol{v}_y) = (B-D)V - \sum_k Q_k f_k(\boldsymbol{y},t) \tag{3.7}$$

where $f_k(\boldsymbol{y},t)$ is the particle distribution in the k^{th} stream, Q_k is the volumetric flow rate of the k^{th} stream, V is the process volume, and $\boldsymbol{v}_y$ is the vector of particle velocities along the $\boldsymbol{y}$ coordinates.

Equation 3.6 is the deterministic form of the general population balance. However, note that just as conservation of mass equations are derived under a continuum assumption concerning the presence of a sufficient number of molecules, the deterministic PBE defines an average value that is only appropriate when the number of particles in the population is sufficiently large. That the PBE reflects an average implies the existence of an underlying probability distribution and the presence of variations or fluctuations about the average value. For situations in which these fluctuations are expected to be of the same order as the average itself, a stochastic formulation is preferred. Section 3.1.3 and Appendix B examine the related field of stochastic population modeling.

3.1.2 Modeling Crystal Size Distribution

The describing variable of interest for crystallization systems is generally a single characteristic crystal size L. Thus, the population balance describes the state of the crystal size distribution (CSD). L is usually defined as either the magnitude of the second largest dimension of the crystals or the radius of a sphere of equivalent crystal volume. The particular choice is often made to suit the experimental sensor technology that is used collect information about the CSD. Since this study uses crystals that are approximately parallel-piped in shape, it is more natural to define L as the second dimension (the width) of the crystals. The motivation for size distribution modeling was presented earlier in Chapter 1.

In addition to crystal size, other variables have been chosen to characterize crystallization processes for the single internal variable case. Many authors have developed models to describe crystallization processes in which the agglomeration of crystals is significant [71, 67, 73, 92]. Though such systems may be equivalently modeled by crystal size, it is often more straightforward to choose particle mass or volume to characterize the distribution. This is because mass and volume, not length, are preserved as particles agglomerate. Saleeby and Lee [133, 85, 134] develop a series of analytical solutions for steady-state crystal size and volume distributions in continuous crystallizers with significant agglomeration kinetics. They examine the effects of size-dependent growth and random growth rate dispersion. Hulbert and Katz [71] characterize particles by mass and develop moment formulations for a simple agglomeration model.

Other investigators have used multiple internal variables. Ilievski and Hounslow [73] use an analytical solution of a two-dimensional steady-state population balance to examine crystal agglomeration in the $Al(OH)_3$ system. The describing variables are crystal volume and zinc tracer content. Zinc tracer seeds replace the standard seed input to a continuous crystallizer operating at steady state. Sieve data are collected over time and analyzed to provide measurements of the volume distribution and zinc content in each size class. The authors fit the data with parameters in two proposed agglomeration models. Though reasonable for a pure agglomeration system, the assumptions used to attain an analytical solution are strict (negligible crystal growth and nucleation, mono-sized seeds, steady-state distribution) and reduce the applicability of the scheme to a small class of problems. However, these restrictions maybe removed if one relies on a stochastic solution method. Gooch and Hounslow [57] propose and solve a variety of multi-variable crystallization problems via a Monte Carlo approach. These researchers use crystal size, growth rate, and number of agglomerated crystallites per crystal, in various combinations, as characteristic variables and demonstrate the flexibility of the Monte Carlo approach. Unfortunately, the restrictions for applying multi-variable approaches are typically the limitations of the measurement technologies used in crystallization processes. It is presently difficult to attain sufficient data to identify the process dynamics that shape the CSD; resolving the kinetics describing distributions in additional variables may prove substantially more difficult.

The reduction of the general PBE to suit the stirred batch cooling crystallization system of this study begins with Equation 3.7. Consider the following assumptions:

- The densitometer sample stream is filtered and assumed solids-free such that $f_{\text{in}} = f_{\text{out}} = 0$. This is the only process fluid stream in the system.
- Volume changes in the system are assumed to be negligible.
- Crystal agglomeration or breakage phenomena are neglected.
- Nucleated crystals appear at a single small size, L_0.

Under the first two assumptions, the stream term in Equation 3.7 is neglected and the equation is put on a constant volume basis. This gives a general equation for the CSD of batch crystallization systems

$$\frac{\partial f(L,t)}{\partial t} = -\frac{\partial (f(L,t)G(L,t))}{\partial L} + (B - D) \tag{3.8}$$

where $\boldsymbol{v}_y$ is replaced by the size-dependent crystal growth rate, $dL/dt = G(L,t)$. At this point, the PBE may be defined for a subvolume of the crystallizer since all subvolumes have the same CSD under the ideal mixing assumption. Therefore, $f(L,t)$, B, and D are redefined on a per unit slurry volume basis. The PBE is appropriate as long as a large number of crystals remains in the smaller volume (see discussion in Section 3.1.3).

Crystal agglomeration and breakage are neglected to simplify the birth and death terms. In agglomeration, small crystals combine into a single larger crystal resulting in "deaths" at small sizes and a "birth" at some larger size. The reverse is true for breakage events. Neglecting these processes limits the birth rate to the nucleation of very small particles and sets the death rate to zero. The nucleation rate is defined as a time-dependent function with strength $B^\circ(t)$ applied at size L_0. Under these assumptions, the batch crystallization form of the general population balance is

$$\frac{\partial f(L,t)}{\partial t} = -\frac{\partial(f(L,t)G(L,t))}{\partial L} + B^\circ(t)\delta(L-L_0) \tag{3.9}$$

where $\delta(L-L_0)$ is the Dirac delta function acting at L_0.

If $L_0 \Longrightarrow 0$, Equation 3.9 may be integrated over the interval L_0^- to L_0^+ to show that the nucleation term may be specified as a boundary condition at $L=0$. The final form of the deterministic PBE for the batch crystallizer is

$$\frac{\partial f(L,t)}{\partial t} = -\frac{\partial(f(L,t)G(L,t))}{\partial L} \tag{3.10}$$

subject to the spatial boundary condition

$$f(0,t) = \frac{B^\circ}{G|_{L=0}}. \tag{3.11}$$

3.1.3 Stochastic Population Models

The PBE is a macroscopic approach that describes the time-averaged behavior of a process that is actually driven by a series of unobservable, microscopic events. In crystallization these events may include the addition of a solute molecule into the lattice of a crystal or a collision between two crystals that results in either an agglomerated crystal or a breakage event. Rather than pursue the time-average explicitly, stochastic models are concerned with the random nature of these events. Their use involves the construction of probability functions and expectations to describe how the particular types of events are likely to affect individual crystals. The macroscopic behavior of the system is predicted by averaging over the trajectories of the individual crystals. For realistic situations, stochastic models are rarely solved analytically but are instead used as the basis of Monte Carlo simulation techniques.

Monte Carlo techniques model a population by following the trajectories of individual particles. In Monte Carlo simulations of crystallization processes crystal growth is generally taken to be deterministic while processes like agglomeration and breakage are considered as jump processes. A jump process is an event that may produce random discontinuities in the state of a particle. A simulation is conducted by allowing the deterministic growth to convect a population of particles smoothly along the size axis for a given period of time. This "quiescent" period is followed by a jump event or a birth or death event. If there is a birth event (nucleation), then a crystal is added to the population (vice versa for a wash-out event). If the event is an agglomeration of two crystals in the population, then two crystals are randomly selected according to a prescribed probability density and one undergoes a jump event while the other is removed from the population. The new coordinates of the crystal experiencing the jump are chosen to preserve the mass of the system (agglomeration preserves mass and volume but not number or size). After the state of the population has been updated, the simulation proceeds with another quiescent period. The lengths of the quiescent periods are also random variables that depend on the size of the population and the rates of the jump processes [137]. The population balance is constructed from these simulations by averaging the state space locations of a large ensemble of individual particle trajectories.

Monte Carlo methods are attractive because, relative to PBE formulations, they are intuitive and easier formulate numerically. They also may be used to extend population models to additional characteristic variables simply by adding new probability densities. This illustrated in the recent work of Gooch and Hounslow [57]. However, Monte Carlo techniques are limited in two respects. First, they are very expensive computationally relative to PBE solutions and may be unwieldy for parameter estimation studies. Naturally, this problem may recede as computers grow progressively faster. Second, few published works have assessed the theoretical equivalence of Monte Carlo simulation and PBE modeling. Ramkrishna [115] professes to make this connection but the relationship is not clear from his presentation. Usually, the equivalence is instead shown via simulations on simple systems [136, 119].

In most instances, the macroscopic behavior of a system is adequately described by the PBE. This is generally true if the modeled population is large because microscopic events tend to occur on short time scales relative to those required to obtain a significant change in the macroscopic CSD. In effect, the microscopic events establish an equilibrium that corresponds to the macroscopic state of the system. However, if the population is very small, this equilibrium concept breaks down and modeling the random CSD fluctuations may become important. Ramkrishna and Borwanker [117] show that the relative magnitude of the random fluctuations is $(\sqrt{\mu_0})^{-1}$, where μ_0 is the expected number of particles in the system.

Fan and coworkers [68, 69] argue that stochastic balances should be used even for large populations. However, their analyses of CSD transients in continuous crystallizers represent no advance over deterministic solutions so it is not clear why they make this suggestion. Ramkrishna [114, 116], only recommends their use for modeling very small populations. Sterilization in microbial cultures and bubble populations in fluidized beds are cited as two such cases. In solution crystallization, the approach has appeal for highly agglomerative systems. However, for most applications in solution crystallizations the deterministic framework is a suitable choice.

Appendix B presents a more detailed look at stochastic modeling. It focuses on the statistical foundations of the method including the Chapman-Kolmogorov and master density equations.

3.2 Material and Energy Balances

It is clear from Equations 3.10 and 3.11 that the evolution of the CSD is driven by the rates of crystal growth and nucleation, processes that are, in turn, driven by the nonequilibrium state of the system. The nonequilibrium state is measured by the relative supersaturation, S, where

$$S = \frac{C - C_{\mathrm{sat}}(T)}{C_{\mathrm{sat}}(T)}. \tag{3.12}$$

Relative supersaturation is a measure of the chemical potential difference for the solute molecules between the solution and crystalline phases. The relationship between relative supersaturation and chemical potential is derived in Appendix C.

Relative supersaturation is a function of both continuous phase solute concentration and system temperature. Thus, the PBE must be augmented by energy and material balances that track these states. These balances are derived from the equations of change using the same assumptions that were applied during the simplification of the PBE in Section 3.1.2. Appendix D provides the detailed derivations of both the mass and energy balances.

3.2.1 Mass Balance

The mass balance models the solute concentration in continuous phase. Since solute concentration is one of the few states of a crystallization system that may be measured experimentally, the mass balance provides a connection between model and experiment that may be used to identify kinetic models and parameters for specific chemical systems.

The form of the mass balance for the batch cooling crystallizer described in Chapter 2 is

$$\frac{d(V\epsilon C)}{dt} = -3\rho_c k_v V \int_0^\infty G f L^2 dL \tag{3.13}$$

where V is the slurry volume, ϵ is the solids-free fraction of V, C is the solute concentration of the continuous phase, ρ_c is the crystal density, and k_v is the volume shape factor converting L^3 to actual crystal volume. As in the derivation of the PBE, stream terms have been neglected.

Under the assumption that V remains constant during the crystallization, Equation 3.13 may be simplified to

$$\frac{d\hat{C}}{dt} = -3\rho_c k_v h \int_0^\infty G f L^2 dL \tag{3.14}$$

where $\hat{C}$ is the mass of solute per mass of solvent and h is a constant factor converting solvent mass to slurry volume. The conversion is made to facilitate the determination of initial concentrations in experimental work. Note that f remains on a per unit slurry volume basis in Equation 3.14.

The mass balance attributes all transfer of solute mass out of the continuous phase to the growth of crystals. This attribution is clarified by taking the growth rate to be size-independent, then the integral term in Equation 3.14 is equal to

$$G \int_0^\infty f L^2 dL \quad \propto \quad \text{growth rate} \times \text{crystal surface area}$$

and the right hand side of Equation 3.13 is seen to multiply a (crystal length)/time term by a crystal area to determine the time rate of change of crystal volume.

3.2.2 Energy Balance

Solution crystallization processes are controlled by manipulating relative supersaturation through changes in liquid phase concentration or solute solubility. In Equation 3.12, solute solubility is expressed as the temperature-dependent saturation concentration, $C_{\text{sat}}(T)$. Cooling crystallization capitalizes on this temperature dependence by using slurry temperature to manipulate the solute solubility. Thus, the crystallizer model must have the capacity to predict accurately the temperature of the slurry. This is the motivation for including an energy balance in the general crystallizer model.[2]

The simplified form of the energy balance for the batch crystallizer is

$$\rho V c_p \frac{dT}{dt} = -3\Delta H_c \rho_c k_v V \int_0^\infty G f L^2 dL - UA\left(T - T_j(t)\right) \tag{3.15}$$

where c_p is the specific heat capacity of the slurry, ΔH_c is the heat of crystallization, U is the overall heat transfer coefficient, A the surface area available for heat transfer, and $T_j(t)$ the time-dependent temperature of the water in the cooling jacket. ΔH_c is negative for the photochemical-heptane system, indicating that the crystallization is an exothermic process.

[2]The energy balance is required to simulate the crystallizer in the absence of slurry temperature data. However, if measurements of slurry temperature are available for a particular experiment the model may be solved by interpolating the data and the energy balance is not required.

For details concerning the physical properties of the photochemical-heptane system, see Appendix A.

Equation 3.15 shows that the energy in the system is a function of two independent processes: internal heat generation due to crystallization and the removal of heat by the jacket cooling water. The crystallization term is essentially the right-hand side of the mass balance multiplied by the heat of crystallization. The crystallization process is controlled by manipulating the jacket temperature in the second term of Equation 3.15. The transfer coefficient term, UA, may be estimated from temperature data correlating slurry temperature changes to steps in $T_j(t)$.

3.3 Crystallization Kinetic Expressions

The crystallization model is completed by the definition of the kinetic processes that relate the dynamics of the crystal population to the state of the bulk system. This involves the specification of rate expressions for crystal growth, nucleation, and dissolution. Each of these processes is a function of relative supersaturation. Growth and nucleation may only occur if this measure is positive, indicating an energetically favorable transition of the solute into the solid phase, while the reverse is true for dissolution. Clearly, as indicated by the spatial boundary condition in Equation 3.11, the relative rates of nucleation and growth are crucial elements in the evolution of the CSD during a crystallization process.

3.3.1 Crystal Growth

A crystal face is a planar surface that grows as existing steps or kinks on the surface are augmented by the incorporation of new solute molecules. Layers spread progressively across the face as new molecules attach themselves to the accessible and energetically favorable leading edges. Macroscopically observable growth is the vertical migration of the plane that results as subsequent layers cover over their predecessors. A fundamental model of this process must describe both how attachment units (e.g., atoms, molecules, ions) are supplied to the existing layers and how new layers are generated.

Burton, Cabrera, and Frank (BCF) [21] propose the accepted fundamental model of layer generation. These authors suggest that layers are initiated by screw dislocations on the crystal surface, similar to that depicted in Figure 3.2. Such crystal imperfections are common and result in persistent generation of attachment sites due to the spiral shape of the dislocation. The authors develop a kinetic theory of growth relating the curvature of the spiral to the relative supersaturation and the spacing of the successive turns. A simple form of the BCF model predicts a macroscopic growth rate of the form

$$G = \frac{k_g'}{k_{\mathrm{bcf}}} S^2 \tanh\left(\frac{k_{\mathrm{bcf}}}{S}\right) \tag{3.16}$$

where S is the relative supersaturation, k_g' is a diffusion rate constant, and k_{bcf} is a spacing-dependent parameter. The BCF model predicts that growth should go from second order in S for small S to first order in S for large S.

The BCF model is a two-step growth model derived for the specific case of molecular attachment to spiral defect hillocks on the crystal surface. The first step is the diffusion of solute molecules from the bulk solution to the crystal surface and the second is the integration of those molecules into the crystal lattice. Many authors have proposed that crystal growth requires such a two-step process. However, the general interpretation is that diffusion is from the bulk to a stagnant film interface and that the difference between

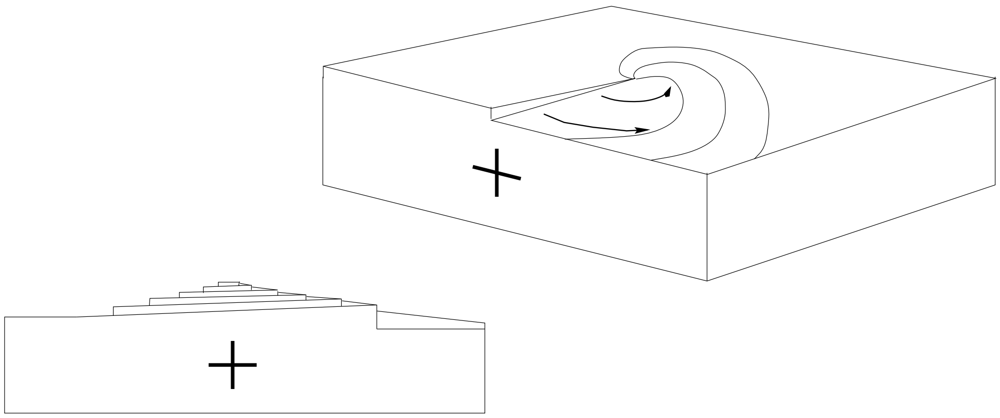

Figure 3.2: Spiral layer formation from a screw dislocation on a crystal surface.

the interface concentration and the saturation concentration drives the subsequent surface integration reaction.

The models of the individual steps may be given as power law expressions

$$\text{bulk diffusion:}\quad G_{bd} = k_{bd}\,(C - C_i)$$

$$\text{surface reaction:}\quad G_{sr} = k_{sr}\,(C_i - C_{\text{sat}})^r$$

where k_{bd} and k_{sr} are rate constants, C_i the solute concentration at the interface, and r the order of the surface reaction. The film, though seldom well-defined in the literature, seems better interpreted as an adsorption layer than as a formal fluid boundary layer. Regardless of the interpretation, the interfacial concentration is unmeasurable resulting in a combined form

$$G = k_{sr}\left(\Delta C - \frac{G}{k_{bd}}\right)^r \tag{3.17}$$

where $\Delta C = C - C_{\text{sat}}$ is the simple supersaturation. If r is assumed equal to 1 or 2, Equation 3.17 may be used to estimate the relative rates of the two-steps.

In most crystallization systems, the surface integration reaction is sufficiently slow relative to diffusion that $C_i \approx C$ and Equation 3.17 reduces to

$$G = k_{sr} \Delta C^r. \tag{3.18}$$

However, many authors [49, 19, 20, 101, 111] have examined alternate cases in which the diffusion step does significantly affect the overall growth rate.

Motivated by the the fact that, like diffusion, crystal dissolution is often a linear function of supersaturation, some authors use an estimate of the dissolution rate constant as an approximation to k_{bd} [101, 28]. However, this method is suspect as crystal shape may change significantly during a dissolution process [111] resulting in different bulk diffusion coefficients in the growth and dissolution regimes. Garside [52] provides a thorough review of two-step growth theories.

This study describes growth as an empirical power law expression in relative supersaturation

$$G = k_g \exp\left(-E_g/RT\right) S^g \tag{3.19}$$

where k_g, g, and E_g are parameters to be estimated for a particular chemical system. Equation 3.19 is the standard expression for growth in the solution crystallization literature and differs from Equation 3.18 in the use of relative supersaturation. The Arrhenius expression is applicable to either expression. Note that the high and low supersaturation asymptotes of the BCF model correspond to $g = 1$ and $g = 2$, respectively.

Equation 3.19 contains no reference to the size of the crystals. Though not a focus of this work, this issue deserves some mention. The size-independent growth rate assumption is often referred to as McCabe's ΔL law. This assumption adequately describes most crystallization systems. However, several authors have observed size-dependent growth rates or have developed models to describe such cases [22, 1, 19, 102]. Garside et al. [53] discuss some possible mechanisms of size-dependent growth. These include bulk diffusion effects due to size-dependent Reynolds number, the Gibbs-Thomson effect relating particle solubility to surface curvature, and defect density on the crystal surface. However, there is no general agreement concerning the actual mechanism.

Rousseau and coworkers [56, 160] argue that apparent size-dependent growth results from growth rate dispersion. This phenomenon describes systems in which individual crystals grow at different rates given identical environmental conditions. Such systems show a gradual spreading of the CSD during growth. Growth rate dispersion has been hypothesized to occur due to either random fluctuations in the growth rates of individual crystals [121] or to the existence of a distribution of constant crystal growth (CCG) rates among the crystal population [77]. The CCG model is supported by the work of Liang et al. [86]. These authors use photomicroscopy to demonstrate CCG rate dispersion for secondary nuclei in a sucrose-water system. In a later, Liang et al. [87] confirm that CCG dispersion also applies to large sucrose seed crystals in batch suspensions.

3.3.2 Crystal Nucleation

As illustrated in Figure 3.3, crystal nucleation is less a well-defined process than an umbrella that covers a diverse array of subprocesses. Nonetheless, nucleation may be broken into the broad categories of primary and secondary nucleation. Primary nucleation occurs in the absence of existing crystals and may occur spontaneously in a pure solution (homogeneous nucleation) or be induced by foreign particles such as dust or impurities (heterogeneous nucleation). Secondary nucleation occurs due to the presence of existing crystals. Classical

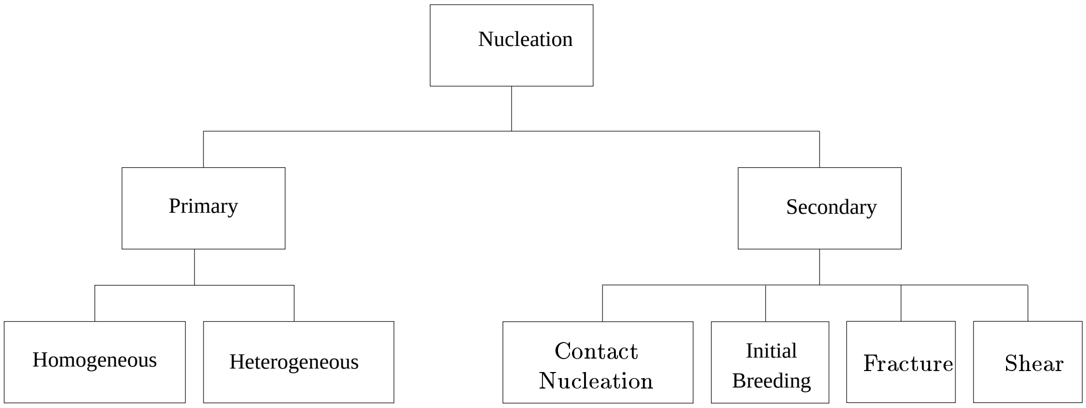

Figure 3.3: Mechanisms of crystal nucleation.

thermodynamic theory and empirical observation both suggest that primary nucleation mechanisms are only initiated by very high levels of supersaturation, well outside the typical operating regimes of solution crystallizers. This fact, and the perpetual presence of large crystal populations in nearly all solution crystallization processes, results in a general neglect of primary nucleation modeling in favor of empirical models of secondary nucleation. This study follows this convention and will, for the remainder of this work, use the term nucleation to apply strictly to secondary mechanisms. However, the reader interested in discussions of primary nucleation may wish to consult the crystallization texts of Mullin [97] and Randolph and Larson [123].

Secondary nucleation occurs to an appreciable degree at relatively low levels of supersaturation and is the dominant form of crystal birth in industrial and lab-scale crystallizers [29]. Effectively, nucleation is a two-part process consisting of the generation of nuclei and their subsequent growth. "Subsequent growth" implies that not all nucleated crystals grow at appreciable rates [77, 51] possibly due to the effects of mechanical stress or lattice strain on or near the face of the nucleated crystals [8, 9]. Some nuclei may even dissolve back into the continuous phase because they are too small to remain stable given the relative supersaturation in the system. This instability is termed the Gibbs-Thomson effect and results from the high surface energies (high surface curvatures) that characterize small particles [46]. Botsaris [11] has listed several potential sources of nuclei, including crystal dendrites, solute clusters near crystal surfaces, and embryos in the supersaturated solution. The author proposes that these sources are converted into active nuclei by fluid shear or by particle-particle or particle-agitator collisions [11, 122].

Of the nucleation mechanisms in Figure 3.3, contact nucleation, which is the result of crystal contacts with agitators, pumps, or other crystals, generally dominates in solution crystallizers. Thus, a large body of work has been dedicated to the study of contact nucleation mechanisms in a variety of chemical systems. McCabe and coworkers [29, 79] perform several types of single crystal experiments with large Magnesium Sulfate crystals. They establish that fluid shear is not a factor for their system and then proceed to examine nuclei yield by contacting a seed crystal with a rod. The rod-crystal contact experiments demonstrate that given constant supersaturation, nuclei production is proportional to contact energy up to a certain point and plateaus thereafter. This result is consistent with the idea that a crystal surface is covered with step edges and other surface structures and that these edges are much more vulnerable to breakage during the contacts than are the planar sections. The yield plateau corresponds to the energy required to break the strongest surface structures. The contact experiments also showed a less well-substantiated proportional relationship between nuclei yield and contact area. A final point of this research concerned the relationship between growth and secondary nucleation. Since the nucleation phenomenon involves the breaking of surface structures, it is necessary that these structures reestablish themselves before a subsequent contact may reproduce the results from an initial contact over the same area. If the structures are reestablished by the process of crystal growth, the regrowth time should correlate to the degree of relative supersaturation in the system. This correlation is confirmed by Johnson et al. [79] for two different supersaturation levels. In each case, the nuclei yield for subsequent contacts increases with the time between contacts until reaching a plateau, signifying a sufficient interval for complete surface regeneration. As expected, higher supersaturation levels result in a smaller regeneration times. These results were later confirmed by Bauer et al. [5] in a similar study of the same chemical system.

McCabe's work provides a basis for the development of macroscopic nucleation models that may be applied to solution crystallization problems. Ottens et al. [104] propose empirical models to describe nucleation due to crystal-agitator and crystal-crystal collisions. Given constant agitation, these models may be written

$$\text{crystal-agitator: } B_a = k_{ba} S^b m_z \tag{3.20a}$$

$$\text{crystal-crystal: } B_c = k_{bc} S^b m_z^2 \tag{3.20b}$$

where m_z is the mass of crystals above a minimum size z. Crystal mass enters the equations due to the assumption that collisional energy is proportional to the momentum of the collision (crystal mass multiplied by the square of the velocity of the collision). The use of a minimum crystal size implies that crystals below a particular size do not participate in contact nucleation. This assumption is reasonable given that smaller crystals are less likely to produce nuclei as they have lower collisional probabilities and energies (small area, small mass, adherence to fluid streamlines). Several researchers [23, 131, 84] provide empirical support for the minimum size assumption. However, the critical size may vary greatly depending on the chemical system. In experiments with a low density Potassium-Alum slurry, Ottens et al. showed nucleation to be a linear function of crystal mass and, in keeping with Equation 3.20a, concluded that crystal-agitator effects were dominant in such a system.

As with the growth expression in Equation 3.19, this study models crystal nucleation as a power law function of relative supersaturation

$$B^\circ = k_b \exp(-E_b/RT) S^b \mu_3^j \tag{3.21}$$

where k_b, b, E_b, and j relate the model to the physical system. Note that the nucleation rate is correlated to μ_3, the third moment of the CSD. Equation 3.21 is a variation on Otten's model, as μ_3 is proportional to the mass of crystals in the crystallizer. Equation 3.21 also may be extended to account for crystals above a critical size through an appropriate change in the calculation of μ_3.

It is a matter of debate whether crystal mass ($\propto$ the third moment of the CSD) or crystal area ($\propto$ the second moment) is more appropriate for use in secondary nucleation models. In recent studies, Wakao et al. [152] and Daudey et al. [33] have estimated parameters for nucleation models based on the second moment. Recall that contact nucleation involves the regeneration of edge structures on crystal surfaces and their subsequent removal by contact. Daudey et al. [33] term these steps surface and mechanical breeding, respectively. They reason that the supply of surface structures is proportional to the total surface area of crystals (i.e., the second moment) while structure removal is a function of collisional intensity (i.e., the third moment). Therefore, systems limited by surface breeding should nucleate in accordance with the second moment and those limited by mechanical breeding should correlate with the third. This suggests that systems with high agitation rates, slow growing crystals, or easily eroded surface structures may be more likely to correlate with the second moment than with the third. However, because the second and third moments are so closely related and model identification involves the simultaneous estimation of a minimum of five parameters (k_g, g, k_b, b, and j), it is unlikely that either model can be clearly discriminated from its rival.

Van der Heijden et al. [149] present an excellent overview of secondary nucleation modeling in which they review many of the empirical results in the literature and specifically addressing issues such as the influence of supersaturation, agitation rate, and CSD.

3.3.3 Dissolution Kinetics

The characterization of crystal dissolution completes the description of the kinetic processes examined in this study. The identification of a dissolution model is useful because it allows consideration of cooling policies that implement a period of negative supersaturation during the crystallization run. Often called a “digestion phase”, a policy of this type may improve the filterability of the slurry by preferentially dissolving very small crystals, due to their higher solubility [156], or by producing crystal shape changes that minimize the tortuousity of the liquid phase path through the filter cake [62].

A standard Fick’s law expression is usually used to model dissolution. However, the order of the process need not be assumed. It is consistent with the structure of the general model to consider dissolution as an extension of the growth and nucleation processes into the negative supersaturation range

$$G = -k_d|S|^d, \qquad S < 0 \tag{3.22a}$$

$$B^\circ = -k'_d|S|^{d'}, \qquad S < 0 \tag{3.22b}$$

where k_d, d, k'_d, and d' are adjustable parameters and the absolute value of S is required to allow dissolution orders other than one. Equation 3.22b is, however, not a realistic physical model because the rate at which dissolving crystals arrive at zero size is also a function of the shape of the CSD. A preferable model would be a deterministic function of the CSD and the dissolution rate of Equation 3.22a. This concept is made clear by imagining the CSD sliding backwards through the zero-size axis at the rate of dissolution. Thus, it may be possible to avoid the estimation of parameters in the modeling of the removal rate. However, such a model requires that the actual shape of the CSD be defined at the start of the dissolution phase.

3.4 Model Summary

This section summarizes the model equations that are implemented in this study. The CSD of the slurry is modeled by a deterministic population balance equation (PBE) which was derived in Sections 3.1.1 and 3.1.2. The PBE model is a partial differential equation in time and crystal size, L, and is given by Equation 3.10. With the additional assumption of a size-independent growth rate, the form as implemented in this study is

$$\frac{\partial f(L,t)}{\partial t} = -G\frac{\partial f(L,t)}{\partial L}. \tag{3.23}$$

The solution phase concentration and slurry temperature within the crystallizer are described by mass and energy balances, respectively. These were derived in Section 3.2 and Appendix D. For size-independent growth rate, they are given as

$$\frac{d\hat{C}}{dt} = -3\rho_c k_v h G \int_0^\infty f(L,t) L^2 dL \tag{3.24}$$

$$\rho V c_p \frac{dT}{dt} = -3\Delta H_c \rho_c k_v V G \int_0^\infty f(L,t) L^2 dL - UA\left(T - T_j(t)\right) \tag{3.25}$$

where $\hat{C}$ is the concentration in the solution phase on a per mass solvent basis, ρ_c is the crystal density, k_v is the volume shape factor, and h converts solvent mass to slurry volume. V is the slurry volume, c_p is the heat capacity of the slurry, and ΔH_c is the heat of crystallization.

In addition to concentration, this study uses transmittance data to estimate the kinetic parameters. The Beer-Lambert law relates the transmittance measurement to the second moment of the CSD.

$$\frac{I}{I_0} = \exp\left(-\frac{l k_a}{2} \int_0^\infty f(L,t) L^2 dL\right) \tag{3.26}$$

I is the intensity of the light transmitted through the slurry, I_0 is the intensity of the incident light, l is the beam length, and k_a is the area shape factor. Details concerning the Beer-Lambert law are given in Section 2.5.3.

Kinetic expressions describing the processes of crystal growth, nucleation, and dissolution were given in Section 3.3. This study does not address temperature-dependent kinetics, therefore only temperature-independent forms of growth and nucleation rate expressions are employed.

$$G = k_g S^g \tag{3.27}$$

$$B^\circ = k_b S^b \mu_3^j \tag{3.28}$$

S is the relative supersaturation. The dissolution model of Equation 3.22a is examined via simulation and experiment but its parameters are not estimated.

The solution of Equations 3.23 through 3.25 require appropriate boundary and initial conditions. The zero-size boundary condition for Equation 3.23 is defined as

$$f(0,t) = \frac{B^\circ}{G|_{L=0}} \tag{3.29}$$

under the assumption that crystals are nucleated at zero-size. The necessary initial conditions for the population, mass, and energy balances, respectively, are

$$f(L,t) = f_s(L), \qquad t = 0 \tag{3.30a}$$

$$\hat{C}(t) = \hat{C}_0, \qquad t = 0 \tag{3.30b}$$

$$T(t) = T_0, \qquad t = 0 \tag{3.30c}$$

where $f_s(L)$ is the size distribution of the seeds.

The model is identified for a specific chemical system by estimating the parameters in Equations 3.27 and 3.28 from experimental data. The parameters to be estimated include k_g, g, k_b, b, and j.

3.5 General Solution Methods

The crystallizer model presented in Section 3.4 consists of a set of coupled, nonlinear integro-differential equations in which the mass transfer of solute into the solid phase is an integral function of the growth of the individual crystals in the population. The population balance is a partial differential equation in time and crystal size. Given realistic models of nucleation and growth, analytical solutions for models of this type are not available and the investigator must resort to numerical techniques. Rawlings et al. [124] and Ramkrishna [116] present recent reviews of numerical solution methods.

PBE solutions in the crystallization literature break down into two broad categories; discretized population balances (DPB) and applications of the method of weighted residuals (MWR). The DPB solution method is essentially a finite difference approximation wherein the length axis of the distribution is sectioned into a finite number of size classes. The population density function is assumed to take some constant average value within each class. The growth rate may be discretized over the size classes, permitting the solution of the PBE for size-dependent growth rate. The DPB solution is popular for modeling systems in which agglomeration or attrition effects are significant. In such cases, the discretization leads to birth and death terms that may be expressed as summations over the discrete size classes [92]. DPB formulations have been proposed by many authors [67, 92, 89]. These variants are distinguished by their choices of discretization intervals and the ways in which they determine the average density function values within each interval. Despite its popularity, the DPB method is not very accurate unless the discretization is very fine. For example, one artifact of this solution method is a systematic over-prediction of the higher moments of the CSD [67, 39]. Recalling the form of the nucleation rate expression of Equation 3.28, it is clear that over-prediction of the third moment will bias the estimation of the kinetic parameters.

Method of weighted residual solutions approximate the population density function as a linear combination of a set of known basis functions, $\Psi_i(L)$,

$$f(L,t) \approx \sum_{i=1}^{N} a_i(t)\Psi_i(L). \tag{3.31}$$

The approximation is substituted into the PBE to define a residual, $r(L,t)$, and that residual is subsequently forced to be orthogonal to a second set of weighting functions, $w_i(L)$, over some domain of L.

$$\int_{L'}^{L''} w_i(L)r(L,t)dL = 0 \tag{3.32}$$

Domains may be chosen as global or as appropriate for a local element solution within a finite element scheme. Galerkin's method describes the case in which the basis and weighting functions are chosen to be the same set of functions. Witkowski and Rawlings [157] use Galerkin's method with Laguerre polynomials on a global domain to describe CSD transients in continuous crystallizers. Other authors have examined similar problems using different types of polynomial functions [26, 155]. The quality of the numerical approximation is generally a function of the order and the degree of similarity between the shapes of the polynomials and the CSD solutions that they are required to represent. Appropriately, Ramkrishna and coworkers [140, 116] recommend the use of problem specific polynomials.

Miller [95] develops an MWR method that solves the PBE using the orthogonal collocation method of Villadson and Stewart [151]. The author uses the technique to compare CSD simulations of a seeded batch cooling crystallizer with and without size-dependent growth rate. Miller cautions that the approximate CSD may contain undesirable oscillations if the shape of the distribution is not easily described by low-order Lagrange polynomials. Should such problems arise, the size domain may be divided into finite elements, adaptive if necessary, over which low-order polynomials may again be appropriate.

3.6 The Method of Moments

Perhaps the most common method used to solve the PBE is the method of moments as developed by Hulbert and Katz [71] and Randolph and Larson [123]. Ramkrishna [140] points out that the method of moments is a special case of the MWR for weighting functions

$$w_i(L) = L^{i-1}. \tag{3.33}$$

This study formulates the PBE assuming that growth is size-independent and that there is no size-dependent destruction of crystals (i.e. no fines dissolution). Under these assumptions the PBE may be stated as

$$\frac{\partial f}{\partial t} = -G\frac{\partial f}{\partial L} + B^{\circ}\delta(L - L_0). \tag{3.34}$$

The ith moment, μ_i, of the CSD is defined as

$$\mu_i = \int_0^{\infty} fL^i dL. \tag{3.35}$$

The rate equation for μ_i is derived by multiplying Equation 3.34 by L^i and integrating over all crystal sizes.

$$\frac{\partial}{\partial t}\int_0^{\infty} fL^i dL = -G\int_0^{\infty} \frac{\partial f}{\partial L}L^i dL + \int_0^{\infty} B^{\circ}\delta(L - L_0)L^i dL \tag{3.36}$$

Performing integration by parts on the first term on the left hand side of Equation 3.36 gives

$$-G\int_0^{\infty} \frac{\partial f}{\partial L}L^i dL = -GL^i f|_0^{\infty} + iG\int_0^{\infty} fL^{i-1}dL. \tag{3.37}$$

If f is finite at $L = 0$ and the distribution contains no crystals of infinite size, the first term on the left hand side of Equation 3.37 vanishes. Substituting Equation 3.37 into

Equation 3.36 and letting $L_0 \Rightarrow 0$, gives

$$\frac{d}{dt}\int_0^\infty fL^i dL = iG\int_0^\infty fL^{i-1}dL + B^\circ 0^i \tag{3.38}$$

or, in terms of the moments

$$\frac{d\mu_0}{dt} = B^\circ \qquad \text{for i} = 0 \tag{3.39a}$$

$$\frac{d\mu_i}{dt} = iG\mu_{i-1} \qquad \text{for i} = 1,2,3 \tag{3.39b}$$

Thus, the PBE is reduced to a set of coupled, nonlinear ordinary differential equations (ODE's). The general crystallizer model summarized in Section 3.4 only requires that the first four moments be determined.

The moment model has some significant limitations. First, the CSD cannot be easily reconstructed by solving a finite set of moment equations because the inverse problem is ill-conditioned [123]. In other words, a finite set of moments does not describe a unique CSD.[3] This uniqueness issue also applies to other MWR techniques. However, the moments are particularly troublesome because CSD's seldom have shapes similar to those of the moment polynomials given in Equation 3.33. A second disadvantage is the inability to close the model equations when general size-dependent functions are included in the integral terms. This prevents closed form moment descriptions of most systems with fines dissolvers, classifiers, or size-dependent growth models. The same concern prevents moments from describing most forms of growth rate dispersion and all but the simplest agglomeration and breakage kernels.

[3]For example, imagine a population consisting of three crystals with dimensionless lengths 1, 2, and 3. The first three moments of the CSD specify the total number of crystals in the population, 3, the sum of the characteristic lengths, 6, and the sum of the squared lengths, 14. Now, imagine a second triplet with lengths 1, 1, and 4. Clearly, the zeroth and first moments are unchanged, 3 and 6, but the second moment is 18. The two CSD's are not identical but this is not apparent from the first two moments. Therefore, the moment set $\{\mu_0 = 3, \mu_1 = 6\}$ does not define a unique CSD.

At the same time, for many systems the advantages of the moment formulation outweigh its disadvantages. First, the convenient reduction of the PBE to a closed set of four coupled ODE's allows the model to be solved quickly, efficiently, and to a specified accuracy by general ODE solution codes. This study uses the differential-algebraic equation solver DASSL [108]. Second, if the CSD cannot be measured quantitatively its calculation, while desirable, is not strictly necessary. Third, because the moments have physical meanings

$$\begin{aligned}
\mu_0 &= \text{total crystal number} \\
\mu_1 &= \text{total crystal length} \\
\mu_2 &\propto \text{total crystal surface area} \\
\mu_3 &\propto \text{total crystal volume}
\end{aligned}$$

measurable quantities may often be represented in terms of the moments. Recall that the transmittance and concentration measurements described in Chapter 2 are functions of μ_2 and μ_3, respectively. To utilize these types of data, other solution techniques must approximate the moments. Fourth, the moments are additive quantities. This is convenient for seeded batch crystallizers that tend to have bi-modal distributions, like that illustrated in Figure 3.4. Moment equations may be formulated for each class and the solutions for the corresponding moments added to provide the total moments for the entire CSD. Lastly, should size-dependent phenomena be considered significant, moment-compatible models are available to allow preliminary investigations of growth rate dispersion [121], size-dependent growth [71], crystal breakage [43], and agglomeration [71, 70].

The moment formulation is a sufficient solution method for this study because the physical apparatus is relatively simple (no fines dissolving or classifiers), the CSD cannot be measured quantitatively, and the measurements are moment-based. The moment model

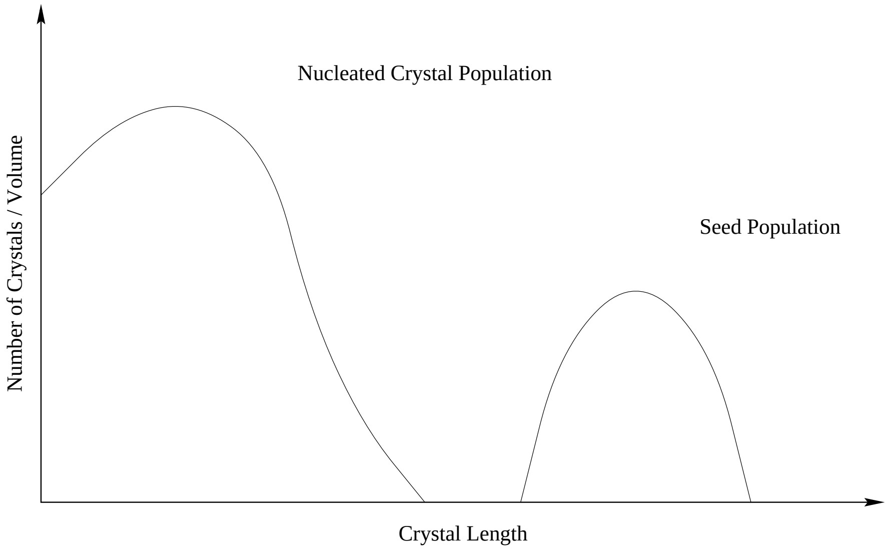

Figure 3.4: Bi-modal crystal size distribution, similar to those often found in seeded batch crystallizers.

is implemented by determining separately the moments of the seed and nuclei classes of the CSD. This strategy is outlined in the following section.

3.7 Model Implementation

This study takes advantage of the additive property of the moments to describe seeded crystallization experiments. Crystallizations of this type tend to produce bi-modal size distributions (see Figure 3.4) characterized by separate classes consisting of the seed and nucleated crystals, respectively. Separate sets of moment equations are used to track the first four moments of each of these classes. The nth CSD moment may be constructed by summing the nth moments of the seed and nucleated crystal classes.

The first four moments of the nucleated class are determined according to Equations 3.39a and 3.39b. For the seed class, the first, second and third moments are determined according to Equation 3.39b. However, because crystal agglomeration and breakage processes are neglected, no new crystals may enter or leave the seed class. Consequently, the zeroth moment of the seed is a constant determined by the initial seed size distribution. This two-class solution technique reduces the partial differential PBE of Equation 3.23 to a set of seven ODE's,

$$\frac{d\mu_{n,0}}{dt} = B^\circ \tag{3.40a}$$

$$\frac{d\mu_{n,i}}{dt} = iG\mu_{n,i-1} \qquad \text{for i} = 1,2,3 \tag{3.40b}$$

$$\frac{d\mu_{s,i}}{dt} = iG\mu_{s,i-1} \qquad \text{for i} = 1,2,3 \tag{3.40c}$$

where the $\mu_{n,i}$ are the moments of the nucleated class and the $\mu_{s,i}$ are the moments of the seeds. The overall moments are defined as $\mu_i = \mu_{n,i} + \mu_{s,i}$. The formulation is completed by the specification of the initial conditions for Equations 3.40a through 3.40c.

In this study, a crystallization experiment begins when the seeds are added to a clear, supersaturated solution. Thus, initially there are seeds but no nuclei. Appropriately, the initial conditions for the moments of the nucleated class are $\mu_{n,i} = 0$, for $i = 0, 1, 2, 3$. The initial conditions for the seed moments are dictated by the seed CSD. The seed CSD is function of the seed mass, the volume of the slurry and the size distribution in the reservoir of crystals from which the seeds are drawn.

3.7.1 Seed Size Distribution

Ideally, the CSD of the seeds may be determined quantitatively by a particle sizer. However, when this is not possible alternate methods employing assumptions about the CSD shape may be used. One such method is to draw the seeds from the segregated size classes of a particle sieve. This allows the limits of the seed CSD to be specified by the screen sizes used in the sieve. The distribution within these limits is still unknown but it seems reasonable to assume it to be parabolic (as in Figure 3.4) or uniform within the limits and zero elsewhere. This method is used by Miller [95] but may not be applied in this study because the photochemical crystals cannot be sieved effectively. Instead maximum and minimum seed crystal sizes are determined by viewing the seeds with optical and electron microscopes. Representative optical and scanning electron micrographs (SEM) are provided in Figures 3.5 and 3.6, respectively. Figure 3.6 is a photograph of dry crystals while the crystals illustrated in Figure 3.5 were sampled from the crystallizer immediately after seeding.

It is difficult to draw any conclusions about the seed CSD from such photographs. However, they can provide more qualitative information regarding maximum and minimum size, crystal shape, and crystal quality. From Figures 3.5 and 3.6 and other photos, the maximum crystal width appears to be between 20 and 25 microns. Since there is no distinct minimum size, the lower range of the seed distribution is taken to originate from zero size.

The specification of the distribution within this range is unknown. However, Figure 3.6 suggests that the greater fraction of crystals resides near the lower size limit. Thus, the seed CSD may be well-described by either a decaying exponential distribution or a decreasing linear distribution that intersects the size axis at the maximum length. Such distributions are illustrated, along with a corresponding uniform distribution, in Figure 3.7. Clearly, the magnitudes of the density axis intercepts in Figure 3.7 are proportional to the mass of the seed load.

To initialize the mass balance properly the mass corresponding to the third moment of the seed distribution must equal the mass of the seeds. The relationship between the two

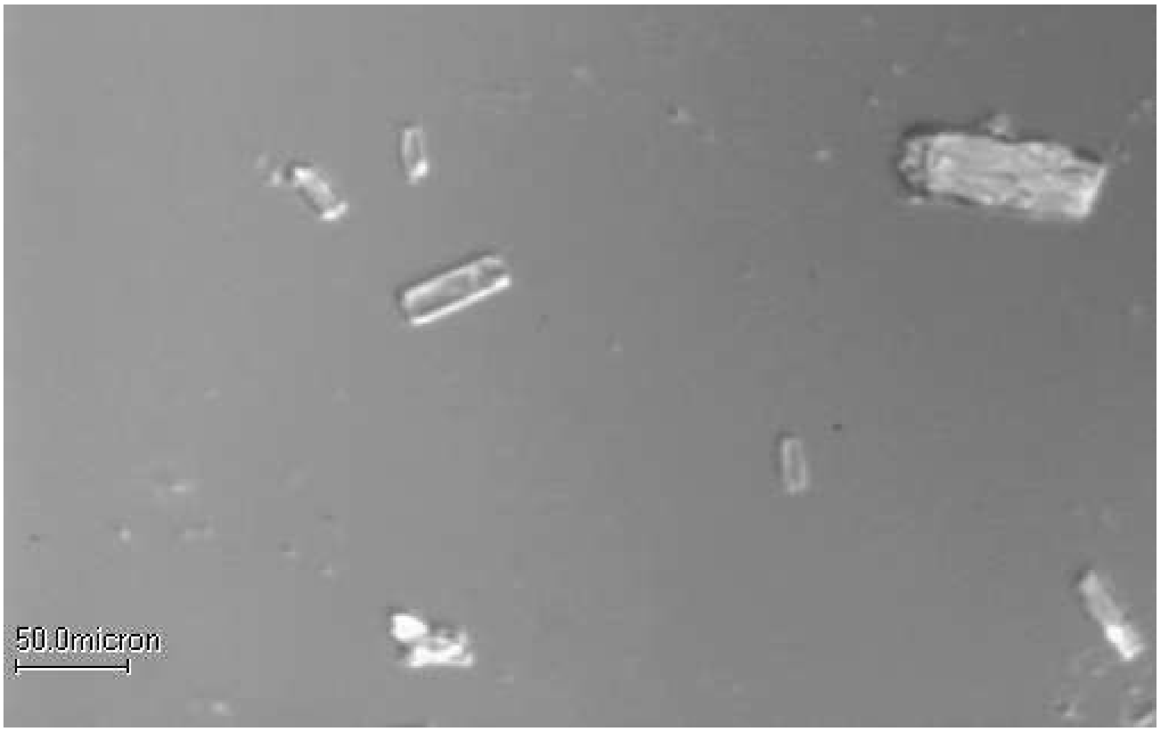

Figure 3.5: Optical micrograph of photochemical seed crystals.

is given as

$$m_c = \rho_c k_v V \int_0^\infty f_s(L) L^3 dL \tag{3.41}$$

where $f_s(L)$ is the seed size distribution. Recalling that the $f(L,t)$ is on a per unit slurry volume basis, take $f_s(L)$ to be proportional to m_c/V for a given seeding. Then Equation 3.41 may be rewritten

$$\left(\frac{m_c}{V}\right) = \rho_c k_v \int_0^\infty \left(\frac{m_c}{V}\right) \overline{f_s}(L) L^3 dL \tag{3.42}$$

where $\overline{f_s}(L)$ may take the forms

$$\overline{f_s}(L) = \begin{cases} f_e \exp(-r_s L), & \text{exponential} \\ f_l \left(1 - \frac{L}{L_m}\right), & \text{linear} \\ f_u, & \text{uniform} \end{cases} \tag{3.43}$$

The linear and uniform distributions are set to zero for $L > L_m$. The constants, f_e, f_l, and

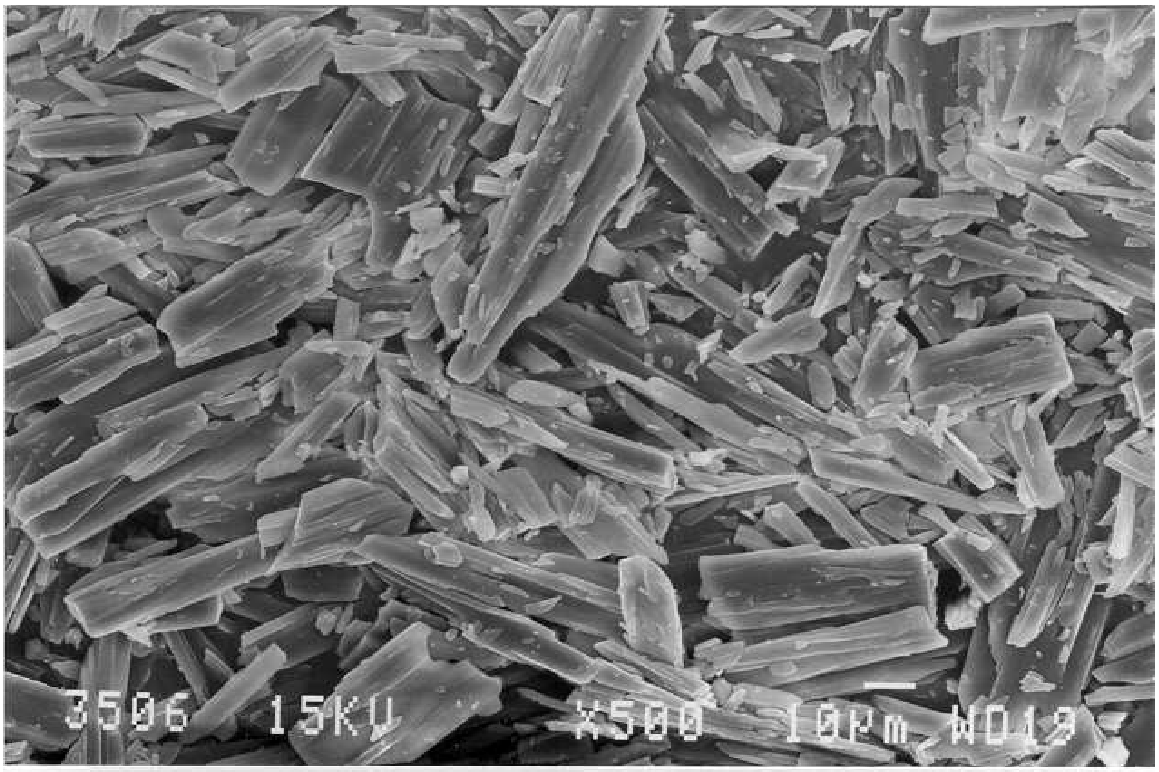

Figure 3.6: Scanning electron micrograph of photochemical seed crystals.

f_u, are fully determined by Equation 3.42 and the seed distribution for any experiment is defined by

$$f_s = \left(\frac{m_c}{V}\right)\overline{f_s}. \tag{3.44}$$

Once a distribution is chosen, the initial conditions for the seed moments may be calculated as

$$\mu_{s,i}(0) = \left(\frac{m_c}{V}\right)\int_0^\infty \overline{f_s}L^i dL. \tag{3.45}$$

Unless otherwise noted, the seed distribution is assumed to be linear and L_m is set 20μm.

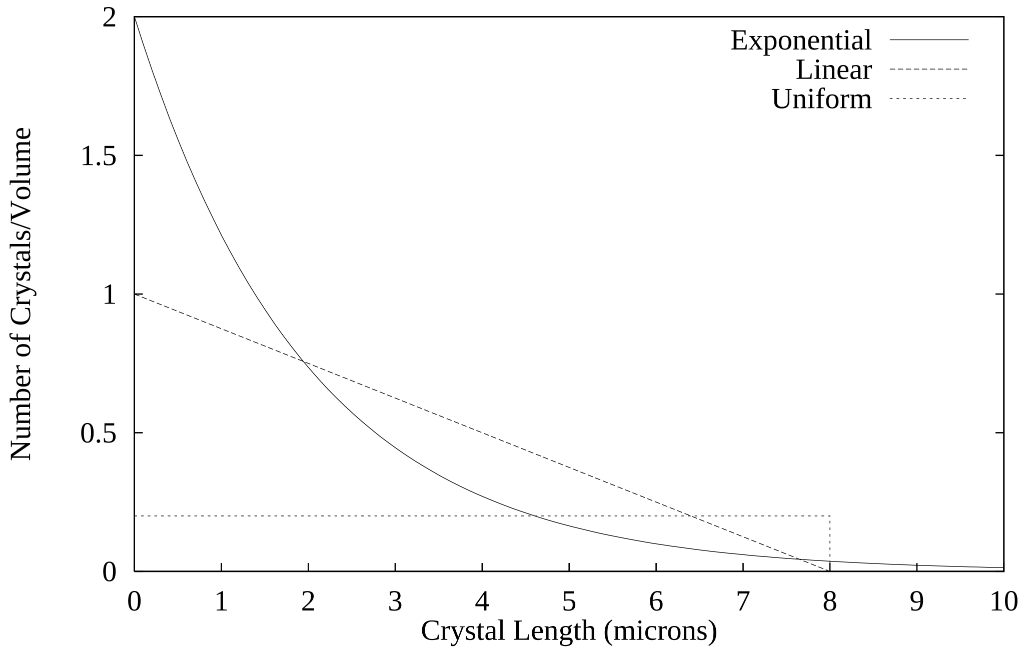

Figure 3.7: Normalized linear, exponential, and uniform seed crystal size distributions corresponding to a given seed mass and slurry volume. Assumes $L_m = 8$ μm for the linear and uniform cases and $r = 0.5$ for the exponential.

3.7.2 Simulations

The definition of the PBE solution method and initial conditions in the preceding sections allows a brief simulation study of the batch crystallizer model. This brief simulation study is designed to provide the reader with a sense of the interplay between the states of the model, the kinetic processes, and the resulting measurement predictions. In particular, three points are addressed in the following sections. First, the relationships between relative supersaturation and the processes of crystal nucleation and growth are explored. The monotonicity of the growth rate with respect to supersaturation is pointed out as is the strong dependence of the nucleation rate on the mass of crystals in the system. Second, the outputs of the response models are examined. The transmittance and concentration

parameter	value	units
k_g	0.30	μm/min
g	2.0	dimensionless
k_b	2.6×10^{10}	(g solvent/cm^3)^j/min
b	3.0	dimensionless
j	2.0	dimensionless

Table 3.1: Kinetic parameters for model simulations.

initial condition	value	units
concentration	0.110	g photochemical/g heptane
temperature	21.0	degrees Celsius
solvent mass	1554.0	g heptane
seed mass	0.10	g photochemical
seed distribution	linear	

Table 3.2: Initial conditions for model simulations.

responses are shown to provide dynamic information at different times during a crystallization. Lastly, the moment trajectories of the nuclei and seed classes are presented with an emphasis on how their relative values impact the crystallization process and reflect the state of the final-time CSD.

Prior to running simulations, nominal kinetic parameters and initial conditions must be specified. These are summarized in Tables 3.1 and 3.2, respectively. The solubility curve, and shape factors required for the mass balance and the transmittance predictions are given in Appendix A. The driving force in the simulation is generated by a cooling profile that decreases linearly from 21.0°C to 18.0°C over a period of 500 minutes. This decrease in temperature corresponds to a 34% decrease in the equilibrium solubility of the photochemical in the heptane solvent.

Supersaturation and Kinetic Processes.

The key feature of a batch crystallization is the relative supersaturation profile that drives the nucleation and growth processes. The supersaturation and third moment trajectories for the simulated experiment are given Figure 3.8. These profiles exemplify the experiments that are presented in Chapter 4. Corresponding plots of the nucleation and growth rates are provided in Figure 3.9.

The simulation begins with a high level of supersaturation. This is allowed by the wide metastable zone of the photochemical system. Initially, the relative supersaturation starts to increase. This occurs because the available crystal surface area is too small and the growth rate too low for the driving force to be relieved by the transfer of mass into the crystalline phase. The low seed mass also limits the rate of secondary nucleation by supporting a low frequency of crystal collisions. The supersaturation peak at 150 minutes occurs as the decrease in $\hat{C}_{\text{sat}}(T)$ with T is surpassed by the decrease in $\hat{C}(t)$ due to high growth rate and increasing crystal surface area (recall Equation 3.24). Since the growth increases monotonically with relative supersaturation this time also gives the peak in the growth rate. However, as the growth and crystal area have increased, so has the crystal mass (third moment). This leads to a short period in which both mass and relative supersaturation are at significant levels, inducing a dramatic increase in the nucleation rate. The wealth of nucleated crystals (high surface area) results in a high rate of solute transfer into the crystalline phase and a subsequent reduction in driving force as the system approaches equilibrium. At the end of the run, growth becomes slow prolonging the approach to equilibrium.

A comparison of Figures 3.8 and 3.9 clearly reveals the driving forces for the kinetic processes. Crystal growth is an exclusive function of supersaturation and therefore mirrors the driving force levels throughout the run. Nucleation, however, requires both supersaturation and crystal mass to be at appreciable levels. Thus, the rate of nucleation peaks mid-way through the simulation but is limited at the beginning and end by low levels of crystal mass and driving force, respectively.

Predicted Measurements.

The two system responses predicted by the model are transmittance and concentration. These are functions of μ_2 and μ_3, respectively. Their trajectories are given in Figure 3.10.

The striking feature in these profiles is the saturation of the transmittance measurement. Recalling that the meaningful range of the Beer-Lambert law is restricted to transmittance values greater than 0.8, the instrument is effectively saturated after 45 minutes. This early saturation is typical of the experiments in this study and indicates the sensitivity of the measurement to the total crystal surface area represented by the product $k_a\mu_2$ (recall Equation 3.26). This condition is unfortunate because it limits the amount useful of transmittance data that may be collected during an experiment. One way to maximize the amount of useful transmittance data is to use small seed loads. Thus, all of the experiments in this study employ seed masses low enough to allow the collection of initial transmittance data.

In the 45 minutes required to effectively saturate the transmittance sensor, the concentration covers only 0.5% of its eventual range. This compares to a 22% change over a later 45 minute period starting 200 minutes into the run. Thus, the concentration data provide little information concerning early-time dynamics but are more informative later in the experiment. In concert with the previous observations about transmittance, this result establishes that the two measurements serve more as separate sensors of early and late-time dynamics than as simultaneous observers of μ_2 and μ_3.

Seed and Nuclei Modes.

The moments of the CSD are the sums of the corresponding moments of the nuclei and seed classes. Since filtration performance is degraded by size variation and the presence of small crystals (nuclei) [82], gauging the relative magnitudes of these CSD classes may prove useful in the optimization of the solution crystallization process. One way to examine relative magnitudes of the moments from the two classes is to plot the ratio $\mu_{n,i}/\mu_{s,i}$ for all of the moments. These profiles are given in Figure 3.11.

Note that none of the moment ratios attain significant values prior to 100 minutes. This reflects the initial suppression of the nucleation rate depicted in Figure 3.9 and points out that the early-time μ_2 dynamics captured by the transmittance data are due entirely to seed growth. As the nucleation rate accelerates, all of the ratios increase dramatically in the later stages of the simulation.

The small size of the nucleated crystals is verified by examining the final values of the moment ratios. The zeroth moment ratio ($\mu_{n,0}/\mu_{s,0}$) shows that nearly 200 times more crystals have been nucleated than seeded. However, the total area of the nucleated class is only 5 times that of the seeds (second moment ratio) and the mass ratio (third moment ratio) is only 1.4. Having five times the area implies that at the end of the simulation over 80% of the mass being transferred into the crystalline phase is being deposited on nucleated crystals.

The size variation in the population may be examined by calculating the coefficient of variation, c_v. The coefficient of variation is defined as the ratio of the standard deviation of the CSD to its mean size. In terms of the moments, it is defined as

$$c_v = \left(\frac{\mu_2 \mu_0}{\mu_1^2} - 1 \right)^{1/2}. \tag{3.46}$$

The profiles of the c_v's for the seeds and the entire population are given in Figure 3.12. Since the variance of the seed distribution does not change (no growth rate dispersion), its deviation is constant and its c_v grows smaller as the average seed size grows larger. The flattening of the curve in the second half of the simulation results from both the approach to an asymptotic bound and the progressive decrease in the solute mass deposited on the seed crystals due to the increasingly large number of nuclei. The c_v of the entire size distribution initially tracks that of the seeds before beginning to increase as new crystals enter the distribution through nucleation. After the nucleation peak, the c_v descends again as the nucleation rate drops and the crystals nucleated during the nucleation burst begin to grow out.

The simulation data show that the nucleation process has a significant effect on the CSD. However, by modeling both the nucleation and seed classes of the distribution the batch crystallizer model may assess the relative magnitude of this effect. In Chapter 5, this capacity is utilized to determine optimal control profiles that minimize the mass of nucleated crystal relative to the mass of the seed class. These profiles are shown to result in improved filtration characteristics for the final-time slurries.

3.8 Summary

This chapter presented the development of the general batch crystallizer model used in this study. The model describes the dynamic evolution of the crystal size distribution using a population balance equation that is solved via the method of moments. Mass and energy balances were derived to describe the solute concentration and system temperature, respectively. The mass and energy balances are related to the PBE through the second moment, μ_2.

Solute concentration and slurry temperature are required to define the relative supersaturation driving force for the kinetic processes of crystal nucleation and growth. Thus, empirical expressions relating nucleation and growth to supersaturation were also presented. In order for the general model to describe the photochemical-heptane system, the parameters in these kinetic models must be estimated from experimental data. The precise estimation of these parameters is the focus of Chapter 4.

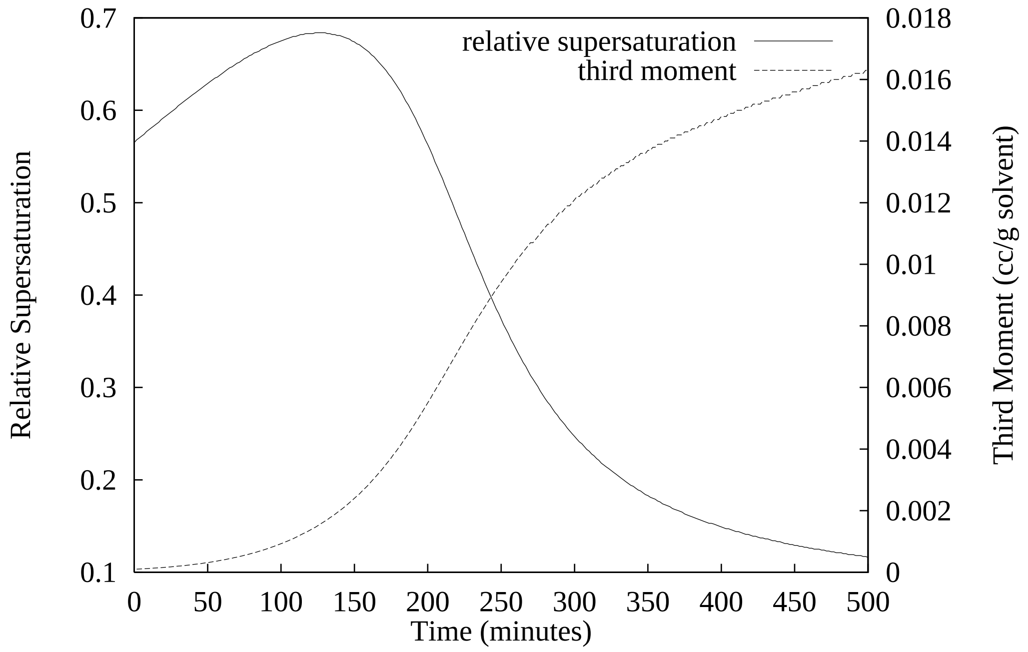

Figure 3.8: Trajectories of relative supersaturation and the third moment of the CSD.

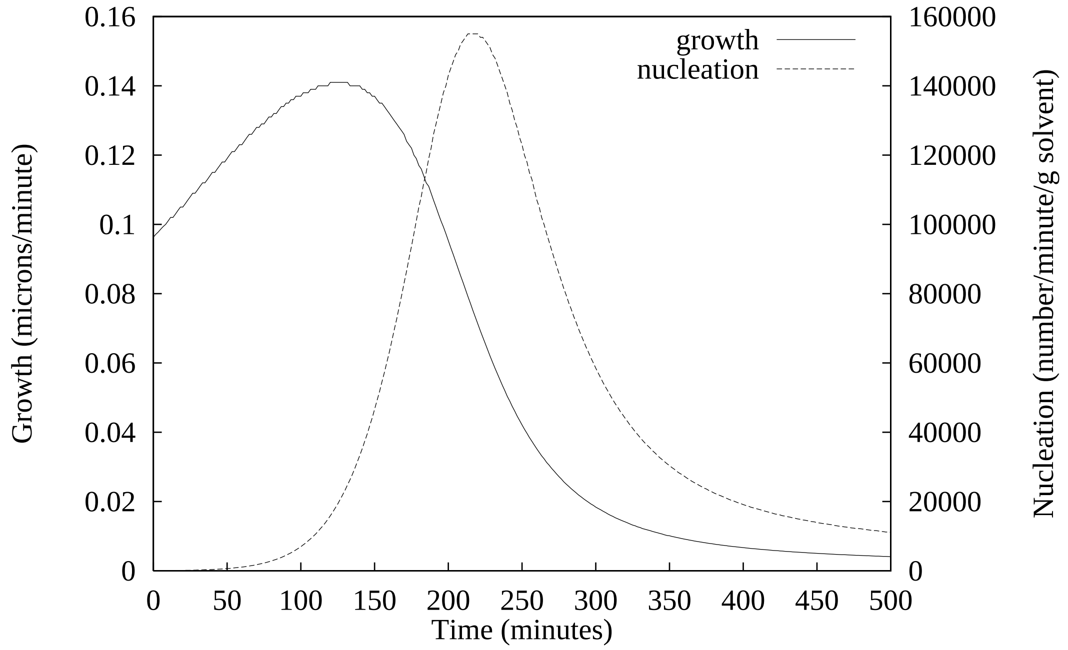

Figure 3.9: Rates of crystal growth and nucleation.

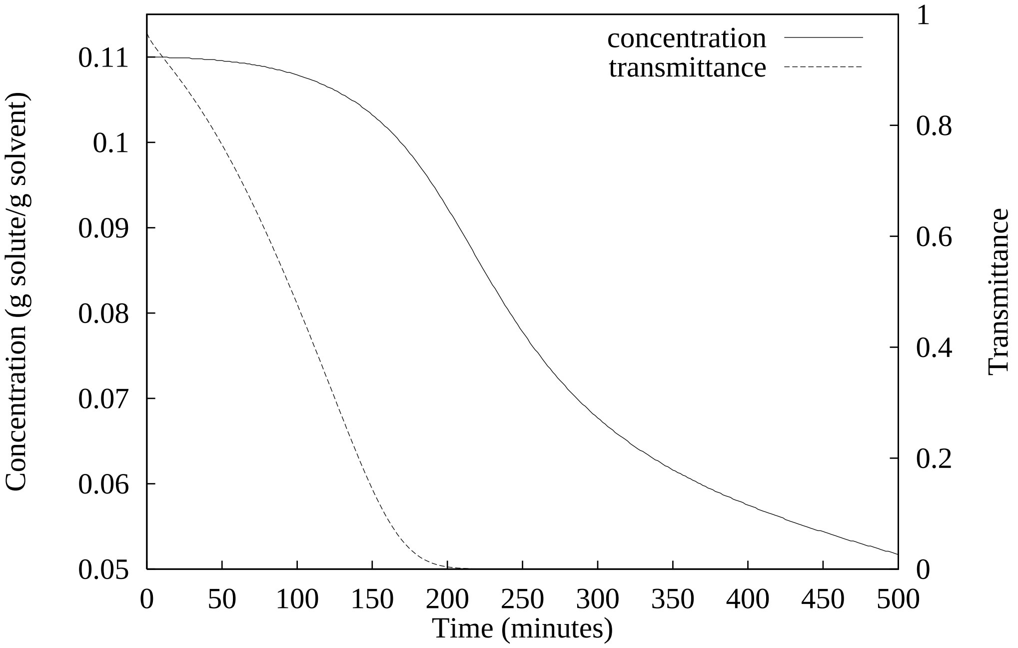

Figure 3.10: Simulated transmittance and concentration profiles.

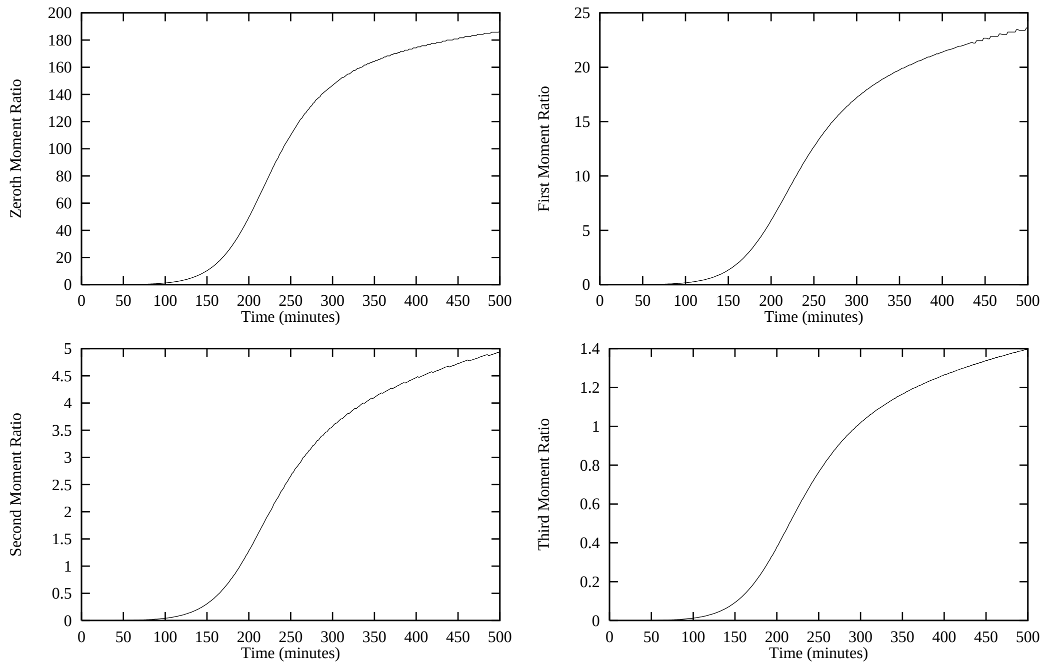

Figure 3.11: Trajectories of the ratios, $\mu_{n,i}/\mu_{s,i}$ for $i = 0, 1, 2, 3$.

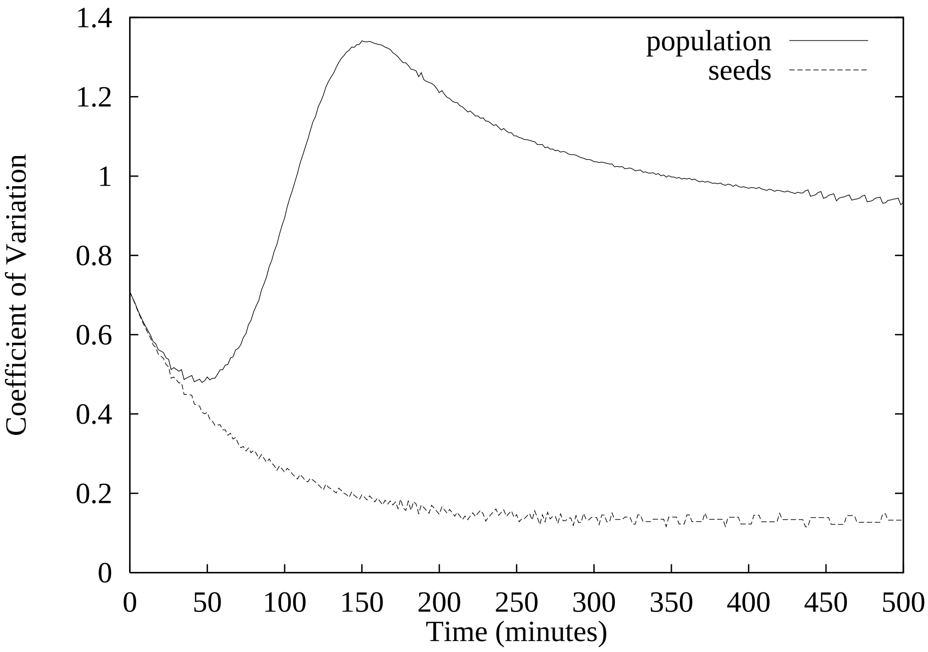

Figure 3.12: Coefficient of variation profiles for the population and seed distributions. Oscillations are due to round-off errors.

Chapter 4

Estimation of Model Parameters

Chapter 3 presented a general model structure for the batch crystallizer apparatus used in this study. The specification of the general model comprises the first phase of a model identification study. The second phase involves the inference of the adjustable parameters in the general model from experimental data collected for a specific chemical system. In this study, the parameters of interest are the kinetic parameters that define the processes of crystal nucleation and growth. These processes drive the evolution of the crystal size distribution during a crystallization. The kinetic parameters must be well-characterized if CSD-dependent performance goals, such as improved filtration, are to be optimized for a particular system.

Parameter estimation is the process of determining the values of the parameters in a general model that allow that model to describe a set of data in an optimal way. Given the experimental conditions necessary to define an experiment and initialize the model, a data set, $\boldsymbol{Y}$, is compared to a corresponding set of model predictions, $\tilde{\boldsymbol{Y}}(\boldsymbol{\theta})$, that depends on a set of model parameters, $\boldsymbol{\theta}$. The comparison produces a set of errors defined by the equation

$$\boldsymbol{Y} = \tilde{\boldsymbol{Y}}(\boldsymbol{\theta}) + \boldsymbol{e} \tag{4.1}$$

where $\boldsymbol{e}$ is the error matrix. Parameter estimation is the process of determining the elements of $\boldsymbol{\theta}$ that minimize some measure of the elements of $\boldsymbol{e}$. If the general model is adequate, the measure of $\boldsymbol{e}$ remains small as experimental data spanning the desired operating region of the process are added to $\boldsymbol{Y}$. Small $\boldsymbol{e}$ implies that the optimal set of parameters, $\hat{\boldsymbol{\theta}}$, is adequate to allow the general model to describe the physical process over the region of interest.

The description given above does not comment on the parsimony of the parameter set, the uncertainty of the parameters, or the sequential revision of the model. However, in concert with Figure 4.1 it does provide a feel for role of the parameter estimation step in a model identification study. Parameter estimation compares the model to the data. If the fit is unacceptable revisions are made in either the model or the apparatus, new experiments are designed, new data collected, and the estimation step is repeated. The process continues until the investigator is satisfied that the identified model is sufficient to act as a surrogate for the physical system in subsequent work (e.g. simulation, control, optimization). This chapter presents the identification process for the kinetic parameters in the batch crystallization model of Chapter 3.

Chapter Outline. The chapter begins in Section 4.1 with a review of previous efforts in the field of kinetic parameter estimation for crystallization systems. Section 4.2 presents a probabilistic approach to parameter estimation that is focused on the use of Bayesian techniques. Independence assumptions and linear uncertainty measures are described and

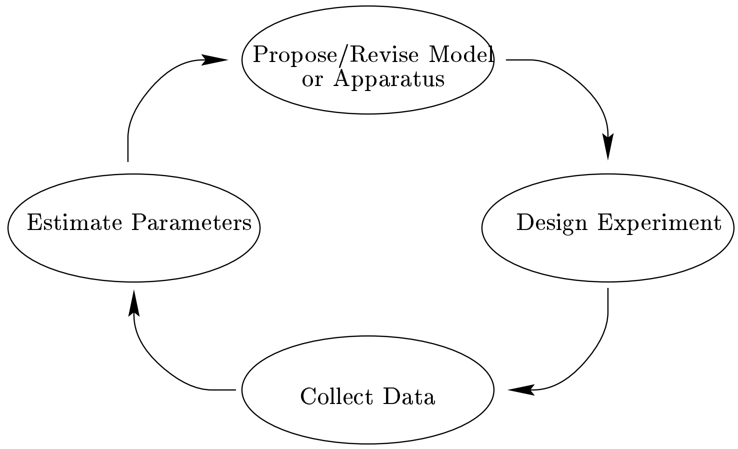

Figure 4.1: The iterative process of model identification.

a nonlinear, Bayesian objective function is defined for the optimization-based estimation method described in Section 4.3. Initial experimental designs and parameter estimates are presented Sections 4.4 and 4.5. A significant portion of Section 4.5 discusses the identification and correction of a bias in the data. Section 4.6 presents two model revisions that were used to allow the general model structure to fit simultaneously the concentration and transmittance data used in the estimation effort. These revisions include a dynamic model of crystal habit and a size-dependent model of nucleation. Section 4.7 presents the improved parameter estimates corresponding to the revised model. Section 4.8 introduces the topic of optimal experimental design, describes the determination of D-Optimal and A-Optimal designs using the revised parameter set, and illustrates a reduction in parameter uncertainty when experimental data for an A-Optimal design are added to the data set. Section 4.9 discusses the use of optical microscopy to confirm the habit changes predicted by the model revision of Section 4.6. Section 4.10 examines some data incompatibility issues spawned by the data bias identified in Section 4.5 and introduces a disturbance model to account for remaining biases. Section 4.12 examines serial correlation in the prediction errors and the adequacy of the linear Hessian for the calculation of parameter confidence intervals.

4.1 Parameter Estimation in Solution Crystallization.

The parameter estimation techniques that have been applied to solution crystallization problems divide neatly into two camps. Early researchers, largely motivated by the lack of mathematical complexity, used the analytical steady state solution to the mixed suspension, mixed product removal (MSMPR) continuous crystallizer to determine graphically parameters for the nucleation and growth processes. A brief discussion of this field is given in Section 4.1.1. Later, as the numerical solution of the nonlinear dynamic PBE became more tractable, a shift was made to the use of dynamic data from MSMPR transients and batch crystallizers. This shift involved a move from graphical estimation methods to the use of optimization based parameter estimation schemes. These latter efforts, reviewed in Section 4.1.2, are similar to that implemented in this study.

4.1.1 Estimation with Steady State MSMPR Crystallizers.

Consider the mixed suspension, mixed product removal (MSMPR) continuous crystallizer form of Equation 3.7 with size-independent growth and a solids-free feed stream.

$$\frac{\partial f}{\partial t} + G\frac{\partial f}{\partial L} = -\frac{f}{\tau} \tag{4.2}$$

τ is the mean residence time equal to V/Q, where Q is the volume flow rate through the crystallizer. Subject to the zero-size boundary condition of Equation 3.11, this form of the PBE has the steady state solution

$$f_{\text{ss}}(L) = \frac{B^{\circ}_{\text{ss}}}{G_{\text{ss}}} \exp\left(-L/G_{\text{ss}}\tau\right) \tag{4.3}$$

where the subscripts indicate steady state values. Given experimental measurements of the steady state CSD, typically provided by sieving slurry samples from the outlet stream, the linear relationship between $\ln(f)$ and L is used to estimate the growth and nucleation rates from the slope and intercept, respectively. The parameters in the models are estimated by correlating rates determined at different steady state supersaturations and slurry densities. Chen and Larson [27] and Garside and Shah [54] provide reviews of parameter estimation studies that have been based on the steady state PBE.

While the steady state MSMPR method has the advantage that it does not require the solution of the dynamic PBE, it is limited in most other regards. Because the nucleation rate is determined by extrapolating the linear fit to zero-size, it is especially sensitive to noisy data. Since sieve data is inherently noisy due to crystal shape effects, clumping, and the discretization of the size classes, estimates of nucleation rates tend to be uncertain. This problem be addressed by determining the intercept from the slurry density [146, 61]. Nucleation rate determination is further complicated when the assumptions of Equation 4.3 are not met. The assumptions may be incorrect when the mixed-product removal assumption is not met or in systems exhibiting size-dependent growth or growth rate dispersion. In these cases, the linear relationship between $\ln(f)$ and L often breaks down in the small size range.

Throughout the extensive literature on MSMPR studies, there is very little effort to evaluate the uncertainty in the kinetic parameter estimates. Failing to measure uncertainty has led to confusion about the sources of the disagreements among sets of kinetic parameters estimated for the same chemical system by different researchers. Clearly, as discussed at length by Garside and Shah [54] and Jones et al. [81], there may be many substantive reasons for such differences including reactor specific hydrodynamics and mismatched driving force ranges. However, the failure to assess uncertainty in the parameter estimates has allowed the discussion to ascend directly into phenomenological debate without first ascertaining whether or not the rival parameter sets are statistically different. Avoidance of this confusion is one reason why acceptable estimation techniques must provide a measure of parameter uncertainty.

A final criticism of MSMPR methods involves the use of uninformative data. The MSMPR method avoids, by definition, the use of dynamic CSD data. Dynamic data are preferred over steady-state data because they provide a greater range and density of the critical states that determine the rates of growth and nucleation, namely relative supersaturation and slurry density (slurry density is $\propto \mu_3$). In particular, since steady-state values of supersaturation tend to be very small, kinetic parameters determined from such data may not be appropriate for the higher driving forces that occur during start-ups, transients, or batch operations. Although these issues were recognized by Timm and Larson [146] in 1968, the use of dynamic data was limited until the advent of optimization based estimation methods.

4.1.2 Optimization based Parameter Estimation.

Optimization provides an intuitive approach to parameter estimation for nonlinear batch systems. The approach uses a nonlinear optimization routine to generate successive guesses of the kinetic parameters. For each guess, the dynamic equations describing the process are solved and the resulting measurement predictions are compared to the data according to an optimization criterion. Generally, the criterion is a measure of the prediction errors. The optimizer uses search algorithms to locate the parameter values that minimize the criterion. Usually, the model equations are solved explicitly by a nonlinear equation integrator (e.g. DDASSL, LSODE) each time the optimizer provides a parameter guess. However, it is also possible to solve the model and optimize the parameters simultaneously [31].

As noted previously, optimization based estimation methods allow the use of dynamic data. The primary sources of dynamic data are MSMPR transients and batch crystallizers. However, the loss of the simplistic steady-state solution substantially reduces the attraction of MSMPR data. From a practical standpoint, continuous crystallizers are cumbersome experimental systems, requiring a large inventory of raw material to maintain controlled feed conditions. By contrast, batch systems offer a flexible and efficient alternative. Tavare [144] summarizes the advantages of using batch experiments as allowing the study of many operating variables in a short period with minimum development time and investment.

Palwe et al. [106] use batch supersaturation data to compare several parameter estimation methods. They note that a least square error optimization technique recovered estimates comparable to those found using polynomial fitting and initial derivative schemes. However, problems with the convergence of the estimates limit their enthusiasm for the method.

Witkowski et al. [159] successfully implement optimization based estimation of nucleation and growth parameters by minimizing the weighted least squares objective function shown below,

$$\Phi_{\mathrm{wls}} = \sum_{k=1}^{m} w_k \left(\sum_{i=1}^{n_k} e_{ik}(\boldsymbol{\theta})^2 \right) \tag{4.4}$$

where m is the number of different types of measurements, w_k weights all measurements of the kth type, n_k is the number of measurements of the kth type, and $e_{ik}(\boldsymbol{\theta})$ is the prediction error for the ith data point of the kth type. Clearly, for single response data ($m = 1$), Equation 4.4 reduces to simple least squares. Witkowski et al. use on-line concentration and transmittance data from a batch Naphthalene-Toluene system. The error weights are chosen to reflect the relative levels of uncertainty for the two sensors, resulting in the preferential minimization of errors in the more precise concentration measurement. However, the actual choice of weights is arbitrary. The uncertainty in the estimates is quantified by the calculation of linear 95% confidence intervals.

Since Witkowski et al., weighted least squares-type objective functions have become popular for multiresponse estimation in crystallization systems. Qiu and Rasmuson [112] use concentration and final-time CSD data to determine kinetic parameters and linear 95% confidence intervals for an aqueous Succinic Acid system. In a later paper [113], these authors verify their earlier work and provide comparison estimates for different weighting choices. However, they make no statements concerning how the weights should be chosen or in what way the weights relate to the probabilistic structure of the prediction errors. Several other researchers have utilized weighted least squares to estimate kinetics in batch systems [34, 32, 90]. However, none of these authors justify their choices of error weights or provide measures of uncertainty.

Related efforts at Delft University have resulted in improved estimation from MSMPR transients. Jager and coworkers [75, 76] use transient data to identify nucleation, growth, and attrition parameters. These studies use light scattering measurements of crystal volume distributions in an unweighted least squares estimation. However, confidence intervals are not calculated. Eek et al. [37] use a similar transient analysis with a weighted least squares scheme. They weight scattering intensity measurements of transients CSD's in a continuous crystallizer with the measurement variance of the sensor. However, they fail to combine data from their different tests into a simultaneous estimation or to calculate confidence intervals.

The primary weakness of the weighted least squares method is the arbitrary selection of the weights. These weights distort the optimization surface and influence the determination of the optimal parameter set. Since it has been shown that rival sets of crystallization kinetic parameters, each providing good fits to experimental data, may suggest dramatically different optimal operating policies [93], this arbitrariness may have significant consequences in some systems. The probabilistic interpretation of parameter estimation relates the kth weight in Equation 4.4 to the variance in the prediction errors of data collected from the kth sensor. This specification is instructive but because of the varied sources of prediction errors (e.g. model errors, sensor accuracy, sensor precision) it does not necessarily facilitate *a priori* weight selection. It is more helpful to implement estimation methods that provide a consistent structure for minimizing the errors in multiresponse data.

Miller and Rawlings [94] employ a maximum likelihood estimation (MLE) scheme that does not require the investigator to specify weights. The likelihood function is the joint probability density of the prediction errors given the data. As shown in Appendix E, the probabilistic structure of MLE allows determination of *a priori* expressions for the prediction error variances in terms of the optimal parameters. Miller and Rawlings use MLE to estimate kinetic parameters from on-line concentration and transmittance measurements in a KNO_3–H_2O batch system. They report linear 95% confidence intervals and subsequently determine optimal control strategies. The sensitivity of the optimal control profiles to the uncertainty in the parameters is also evaluated. This study applies a probabilistic estimation method similar to that of Miller and Rawlings.

4.2 Probabilistic Interpretation of Parameter Estimation

Probability theory provides a unifying approach for parameter estimation problems. Simply put, given the data $\boldsymbol{Y}$ and the general model, the goal is to find the probability distribution of the model parameters, $\boldsymbol{\theta}$. The optimal parameter estimates are those that locate the maximum, or mode, of the probability distribution. They are sometimes referred to as the *modal* or *point estimates*. Measures of uncertainty are identified by quantifying the spread of the distribution about the mode.

Consider the general problem of estimating the elements of the parameter vector $\boldsymbol{\psi}$ given the model and the data. $\boldsymbol{\psi}$ contains both the model parameters $\boldsymbol{\theta}$ and any other parameters, such as prediction error variances, that may be required to specify fully the distribution of $\boldsymbol{\theta}$. Derived from the definition of conditional probability [6],

$$p(\boldsymbol{\psi}, \boldsymbol{Y}) = p(\boldsymbol{\psi}|\boldsymbol{Y})p(\boldsymbol{Y}) = p(\boldsymbol{Y}|\boldsymbol{\psi})p(\boldsymbol{\psi})$$

Bayes' Theorem provides a formula for the inference of $\boldsymbol{\psi}$ from the data, $\boldsymbol{Y}$.

$$p(\boldsymbol{\psi}|\boldsymbol{Y}) = \frac{p(\boldsymbol{Y}|\boldsymbol{\psi})p(\boldsymbol{\psi})}{p(\boldsymbol{Y})} \tag{4.5}$$

$p(\boldsymbol{\psi}|\boldsymbol{Y})$ is the posterior density of the parameters given the model structure and the data. It summarizes the information about the parameters that is provided both by the data and by any *a priori* knowledge. Prior information regarding the parameters is expressed by the prior distribution, $p(\boldsymbol{\psi})$. Given a data set and a model, $p(\boldsymbol{Y})$ is a constant.

$$p(\boldsymbol{Y}) = \int p(\boldsymbol{Y}|\boldsymbol{\psi})p(\boldsymbol{\psi})d\boldsymbol{\psi}$$

Fisher coined the term "likelihood function" to describe $p(\boldsymbol{Y}|\boldsymbol{\psi})$ when $\boldsymbol{Y}$ is given. The likelihood is written as $\ell(\boldsymbol{\psi}|\boldsymbol{Y})$, giving the form of Bayes theorem most often associated with parameter estimation.

$$p(\boldsymbol{\psi}|\boldsymbol{Y}) \propto \ell(\boldsymbol{\psi}|\boldsymbol{Y})p(\boldsymbol{\psi}) \tag{4.6}$$

$$\text{Posterior} \propto \text{Data} \times \text{Prior}$$

Bayesian estimation has the pleasant interpretation that the parameters are inferred using all previous knowledge plus the new information contained in $\boldsymbol{Y}$. If $p(\boldsymbol{\psi})$ is taken to be a uniform distribution, the posterior distribution is proportional to the likelihood function and Bayesian estimation reduces to maximum likelihood estimation (MLE).

The posterior density function is derived by formulating $\ell(\boldsymbol{\psi}|\boldsymbol{Y})$ and $p(\boldsymbol{\psi})$. These formulations require the specification of distribution and independence assumptions regarding the prediction errors. In this study, these assumptions are as follows:

- The prediction errors are Normally distributed with zero mean and covariance matrix $\boldsymbol{V}$.
- Prediction errors corresponding to different sample times are uncorrelated.
- Prediction errors corresponding to different measurement types are uncorrelated.
- The prediction errors are homoscedastic (error variance is independent of the magnitude of the measured variable).

Bayesian estimation methods are not limited to these assumptions and may be generalized by allowing error correlation or non-Normal distributions.

4.2.1 The Likelihood Function

Given the distribution and independence assumptions above, the likelihood function may be derived from the appropriate form of $p(\boldsymbol{Y}|\boldsymbol{\psi})$. The $p(\boldsymbol{Y}|\boldsymbol{\psi})$ is the joint probability density of the collected data that make up $\boldsymbol{Y}$, given the parameters $\boldsymbol{\psi}$. The model predicts $\boldsymbol{Y}$, subject to the $\boldsymbol{\theta}$ subset of $\boldsymbol{\psi}$

$$y_i = \tilde{y}_i(\boldsymbol{\theta}) + e_i \tag{4.7}$$

where y_i is the ith measurement in $\boldsymbol{Y}$, $\tilde{y}_i(\boldsymbol{\theta})$ the corresponding model prediction, and e_i the corresponding prediction error. With the prediction errors assumed independent, $p(\boldsymbol{Y}|\boldsymbol{\psi})$ is the product of the distributions of the individual errors

$$p(\boldsymbol{Y}|\boldsymbol{\psi}) \approx \prod_{i=1}^{n} p(e_i|\boldsymbol{\psi}) \tag{4.8}$$

where n is the number of measurements in $\boldsymbol{Y}$. Since each error is Normally distributed with zero mean this product has the general form

$$p(\boldsymbol{Y}|\boldsymbol{\theta}, \boldsymbol{\sigma}) = (2\pi)^{-n/2} \left|\boldsymbol{V}\right|^{-1/2} \exp\left(-\frac{1}{2}\boldsymbol{e}^T(\boldsymbol{\theta})\boldsymbol{V}^{-1}\boldsymbol{e}(\boldsymbol{\theta})\right). \tag{4.9}$$

Note that in Equation 4.9, $\boldsymbol{\psi}$ is replaced by $\boldsymbol{\theta}$ and $\boldsymbol{\sigma}$, where $\boldsymbol{\sigma}$ consists of the square roots of the independent elements of $\boldsymbol{V}$. This change denotes that, under the Normality assumption, the distribution of $\boldsymbol{Y}$ is specified entirely by the model parameters, $\boldsymbol{\theta}$, and the elements of the measurement covariance matrix, $\boldsymbol{V}$.

The independence assumptions specify that $\boldsymbol{V}$ is a diagonal matrix containing m independent elements, where m is the number of measured variables employed in the estimation. These assumptions allow Equation 4.9 to be further simplified to

$$p(\boldsymbol{Y}|\boldsymbol{\theta}, \boldsymbol{\sigma}) = \left(\sqrt{2\pi}\right)^{-\left(\sum_{k=1}^{m} n_k\right)} \prod_{k=1}^{m} \sigma_k^{-n_k} \exp\left(-\frac{1}{2}\sum_{k=1}^{m} \sigma_k^{-2} \sum_{i=1}^{n_k} e_{ik}^2(\boldsymbol{\theta})\right) \tag{4.10}$$

where n_k is the number of measurements of the kth type, σ_k is the kth element of $\boldsymbol{\sigma}$, and $e_{ik}(\boldsymbol{\theta})$ is the prediction error for the ith measurement of the kth type. Clearly, $\boldsymbol{\sigma}$ is now an m-dimensional vector. For given $\boldsymbol{Y}$, as opposed to given $\boldsymbol{\theta}$ and $\boldsymbol{\sigma}$, Equation 4.10 is the likelihood function, $\ell(\boldsymbol{\theta}, \boldsymbol{\sigma}|\boldsymbol{Y})$.

Equation 4.10 is the expression used in the MLE work of Miller and Rawlings [94]. The parameters locating the mode of this form of the likelihood function are those that minimize the objective function

$$\Phi_{\text{mle}} = \sum_{k=1}^{m} n_k \ln \left(\sum_{i=1}^{n_k} e_{ik}^2(\boldsymbol{\theta}) \right). \tag{4.11}$$

Given the modal parameter set, $\hat{\boldsymbol{\theta}}$, estimates of the m elements of $\hat{\boldsymbol{\sigma}}$ are determined by

$$\hat{\sigma}_k = \left(\frac{\sum_{i=1}^{n_k} e_{ik}^2(\hat{\boldsymbol{\theta}})}{n_k} \right)^{1/2}. \tag{4.12}$$

Derivations of Equations 4.11 and 4.12 are given in Appendix E. A general discussion of MLE theory and application techniques is provided by Bard [4].

4.2.2 The Prior Distribution

The prior distribution of the parameters expresses the investigator's knowledge of the parameters prior to the collection of data. Priors may be informative or noninformative, depending on whether or not they are intended to bias the shape of the posterior distribution. The choice is often dependent on the existence and perceived quality of the *a priori* information.

Informative priors are designed to influence the shape of the posterior density function based on the *a priori* knowledge. As such, informative priors are the basis of the optimal experimental design methodology described in Section 4.8. In experimental design, the posterior density function determined from an initial set of experiments is used as an informative prior for the design of a subsequent experiment. The new experiment is designed to maximize some measure parameter information in the predicted posterior density function, given the informative prior. Informative priors are most useful in situations of this type in which prior information of acceptable quality exists but is nonetheless deemed insufficient for the investigator to halt the identification process. However, since model identification studies are generally motivated by the lack of existing information, informative priors are rarely available for the study of new systems and noninformative priors are required.

Noninformative priors are designed to let the likelihood function (i.e. the data) shape the posterior distribution of the parameters. They are derived through an analysis of the likelihood function. Jeffreys [78] presents a method for generating a noninformative prior that makes the posterior distribution a *data-translated* distribution. The term data-translated implies that different data sets reposition the mode of the likelihood with respect to the parameter space but that the shape of the distribution is unaffected by such relocations (i.e. the shape is independent of $\boldsymbol{\psi}$). If the shape is not independent of $\boldsymbol{\psi}$, then a particular parameterization, $\boldsymbol{\phi}(\boldsymbol{\psi})$, may be determined for which the independence is preserved. From this reasoning, Jeffrey's derived the following general equation for his prior,

$$p(\psi) \propto \sqrt{|\boldsymbol{I}_n|} \tag{4.13}$$

where $\boldsymbol{I}_n$ is the Fisher information matrix defined as

$$\boldsymbol{I}_n = -E\left(\frac{\partial^2 \ln p(\boldsymbol{Y}|\boldsymbol{\psi})}{\partial\psi_i \partial\psi_j}\right). \tag{4.14}$$

The information matrix is the negative of the expected modal curvature of $p(\boldsymbol{Y}|\boldsymbol{\psi})$. Therefore, the Jeffreys prior casts the independence of shape as the independence of modal curvature.

Given the development of the likelihood function in the previous section, the parameter vector $\boldsymbol{\psi}$ consists of the model parameters, $\boldsymbol{\theta}$, and the m-dimensional measurement standard deviation vector, $\boldsymbol{\sigma}$. It is assumed that $\boldsymbol{\theta}$ and $\boldsymbol{\sigma}$ are independently distributed such that

$$p(\boldsymbol{\psi}) \approx p(\boldsymbol{\theta})p(\boldsymbol{\sigma}). \tag{4.15}$$

The elements of $\boldsymbol{\theta}$ are called *location parameters* because their optimal values are determined from the location of the mode of $p(\boldsymbol{\theta}, \boldsymbol{\sigma}|\boldsymbol{Y})$. The elements of $\boldsymbol{\sigma}$ are called *scale parameters* because their modal values affect the spread of the density function. For location parameters, no reparameterization is necessary as the Jeffreys prior generally suggests a locally uniform distribution.[1]

$$p(\boldsymbol{\theta}) = \text{constant} \tag{4.16}$$

For scale parameters, however, a reparameterization is required. In Appendix E, the following nonuniform prior is derived for the form of the likelihood function given by Equation 4.10.

$$p(\boldsymbol{\sigma}) \propto \prod_{k=1}^{m} \sigma_k^{-1} \tag{4.17}$$

Equations 4.16 and 4.17 are the prior distributions implemented in this study.

Given Equation 4.15 and uniform $p(\boldsymbol{\theta})$, $p(\boldsymbol{\psi})$ is clearly proportional to $p(\boldsymbol{\sigma})$. This is in keeping with the general recommendation of Stewart et al. [141] that $p(\boldsymbol{\psi})$ be taken proportional to the Jeffreys prior for the scale parameters. In addition to the prior discussions

[1] The distribution is defined as local to allow normalization to unit probability.

of Jeffreys [78] and Stewart et al. [141], Box and Tiao [16] provide an extensive coverage of issues pertaining to the philosophy and use of prior distributions in Bayesian estimation.

4.2.3 The Posterior Density

As expressed in Equation 4.6, the posterior density of the parameters given the data is proportional to the product of the likelihood function and the prior distribution of the parameters. Given Equations 4.10 and 4.17, the posterior density of the parameters in the crystallization model conditional on the data is

$$p(\boldsymbol{\theta}, \boldsymbol{\sigma}|\boldsymbol{Y}) \propto \left(\sqrt{2\pi}\right)^{-\left(\sum_{k=1}^{m} n_k\right)} \prod_{k=1}^{m} \sigma_k^{-(n_k+1)} \exp\left(-\frac{1}{2}\sum_{k=1}^{m} \sigma_k^{-2} \sum_{i=1}^{n_k} e_{ik}^2(\boldsymbol{\theta})\right). \tag{4.18}$$

The conditional posterior density of Equation 4.18 is the solution to the Bayesian estimation problem. It expresses the relationship between the parameters and the data, subject to the model. Note that the model now implies both the crystallization model of Chapter 3 as well as the probabilistic structure of the prediction errors and the prior distributions. An example posterior density for a two-parameter, linear system is given in Figure 4.2.

Despite constituting the complete solution to the estimation problem, the posterior distribution seldom is constructed explicitly. Instead, expressions are derived from the posterior distribution to allow the direct determination of the modal parameter estimates and prediction error variances. Taking the natural logarithm[2] of $p(\boldsymbol{\theta}, \boldsymbol{\sigma}|\boldsymbol{Y})$ and following a development similar to that performed for the likelihood function in Appendix E, it may be shown that the modal parameters of $p(\boldsymbol{\theta}, \hat{\boldsymbol{\sigma}}|\boldsymbol{Y})$ are equivalent to those minimizing the

[2] In addition to simplifying the analysis, the logarithm facilitates the numerical solution of the modal expressions by compressing the magnitude of the posterior density, which may exceed machine precision for large data sets.

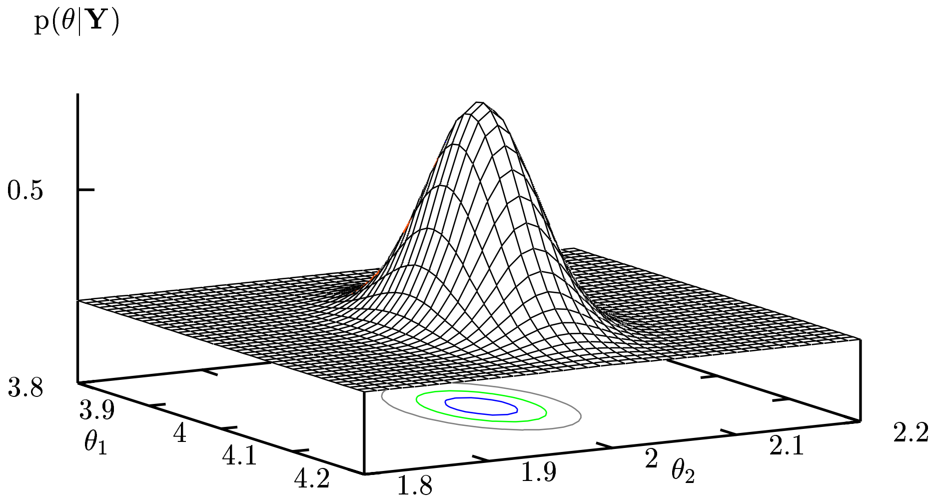

Figure 4.2: Normalized posterior density for a linear, two-parameter system with two measurement types. Ellipses denote lines of constant probability density.

following objective function.

$$\Phi_{\mathrm{be}} = \sum_{k=1}^{m} (n_k + 1) \ln \left(\sum_{i=1}^{n_k} e_{ik}^2(\boldsymbol{\theta}) \right) \tag{4.19}$$

Equation 4.19 corresponds to the following modal estimate of the m elements of $\hat{\boldsymbol{\sigma}}$.

$$\hat{\sigma}_k = \left(\frac{\sum_{i=1}^{n_k} e_{ik}^2(\hat{\boldsymbol{\theta}})}{n_k + 1} \right)^{1/2} \tag{4.20}$$

The posterior distribution corresponding to Φ_{be} is written $p(\boldsymbol{\theta}, \hat{\boldsymbol{\sigma}}|\boldsymbol{Y})$ to denote the use of the modal value of $\boldsymbol{\sigma}$.

From an examination of Figure 4.2, it is clear that the point estimate of the mode (the parameters locating the peak of the distribution surface) given by the minimization of Φ_{be} is an insufficient expression of the information contained in the complete distribution. In particular, point estimates provide no measure of the spread of the posterior distribution about the mode. An approximation of the spread may be obtained by assuming that the model can be represented by linear functions in the vicinity of the modal estimates,

$$\tilde{y}_{ik}(\boldsymbol{\theta}) \approx \tilde{y}_{ik}(\hat{\boldsymbol{\theta}}) + \boldsymbol{B}_{ik}(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}) \tag{4.21}$$

where $\tilde{y}_{ik}(\boldsymbol{\theta})$ is the model prediction for the ith data point of the kth type and $\boldsymbol{B}_{ik}$ is the $1 \times p$ sensitivity vector (where p is the number of parameters) given below.

$$\boldsymbol{B}_{ik} = \left.\frac{\partial \tilde{y}_{ik}}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \tag{4.22}$$

The linearized problem is used to approximate the regions of highest posterior density (HPD) that surround the modal estimate of $\boldsymbol{\theta}$. The boundaries of the HPD regions correspond to contours of constant probability density. When projected into the parameter space (see the base contours of Figure 4.2), each contour is ellipsoidal in shape with a volume that corresponds to some fraction of the total probability expressed by the posterior distribution.

Given the assumptions of this study, the Hessian of the linearized distribution is determined from the expression

$$\boldsymbol{H}_\theta = \sum_{k=1}^{m} \sigma_k^{-2} \sum_{i=1}^{n_k} \boldsymbol{B}_{ik}^T \boldsymbol{B}_{ik}. \tag{4.23}$$

With Normally distributed prediction errors, it may be shown [4] that the value of

$$(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^T \boldsymbol{H}_\theta (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}) \tag{4.24}$$

is distributed as χ^2 with p degrees of freedom. The HPD region constituting $100(1-\alpha)\%$ of the total probability expressed by the linearized posterior distribution is the ellipsoidal region

$$(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^T \boldsymbol{H}_\theta (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}) = \chi_p^2(\alpha). \tag{4.25}$$

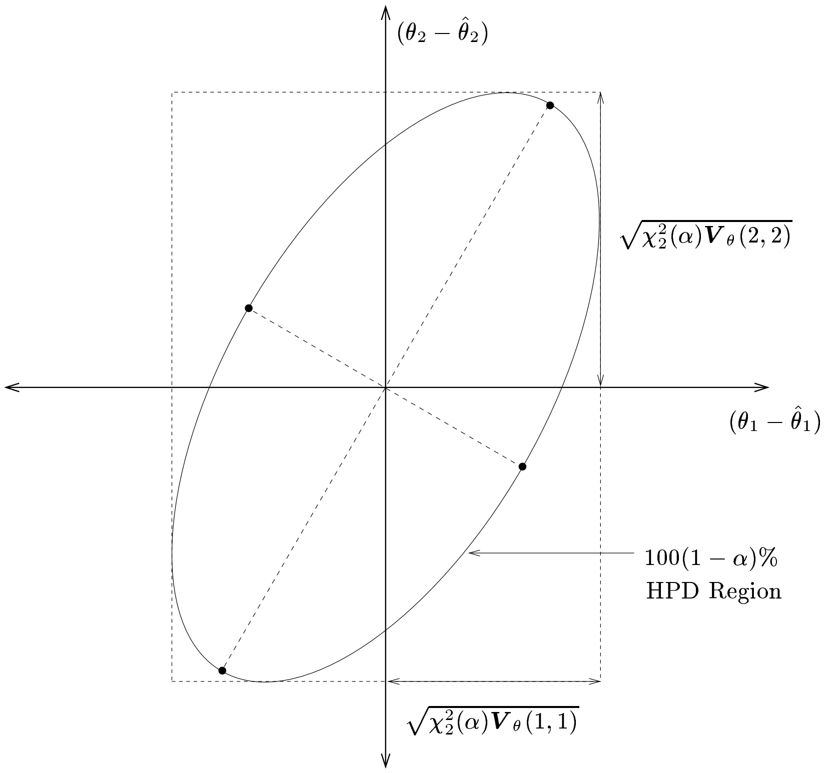

Figure 4.3: Elliptical region of highest posterior density (HPD) for a two parameter system. Conservative confidence intervals are defined by the circumscribing box.

Because this study estimates parameter sets with more than three elements, it is more convenient to obtain conservative estimates of these regions by locating the hyperbox in the parameter space that circumscribes the ellipsoidal confidence region. The dimensions of this hyperbox are determined by

$$\theta_i \approx \hat{\theta}_i \pm \sqrt{\chi_p^2(\alpha)\boldsymbol{V}_\theta(i,i)} \tag{4.26}$$

[41], where $\boldsymbol{V}_\theta(i,i)$ is i,i element of the parameter covariance matrix $\boldsymbol{V}_\theta = \boldsymbol{H}_\theta^{-1}$. The linear 95% confidence intervals cited throughout the remainder of this work were calculated according to Equation 4.26 with $\alpha = 0.025$. A HPD region ellipse and its corresponding confidence interval box is depicted for a two parameter system by Figure 4.3.

For the purposes of this study, modal estimates and linear confidence intervals are used to approximate the information in the posterior distribution.

4.3 Implementation of Bayesian estimation

The kinetic parameters given in Equations 3.27 and 3.28 are estimated from experimental data collected for the photochemical-heptane system. The parameter vector has the form:

$$\boldsymbol{\theta} = \begin{bmatrix} k_g & : & \text{growth rate constant} \\ g & : & \text{growth rate order} \\ k_b & : & \text{nucleation rate constant} \\ b & : & \text{nucleation rate order} \\ j & : & \text{nucleation crystal mass order} \end{bmatrix} \tag{4.27}$$

The parameters are estimated by posing a nonlinear program (NLP) with the Bayesian objective function of Equation 4.19. The objective function is evaluated by a sequential optimization and solution method in which the equations of Section 3.4 are solved explicitly at each iteration of the optimization routine by the differential-algebraic solver DDASSL [108]. The method may be summarized in the following statement of the NLP.

$$\begin{aligned} \min_{\boldsymbol{\theta}} \quad & \Phi_{\text{be}}\left(\boldsymbol{Y}, \tilde{\boldsymbol{Y}}(\boldsymbol{\theta}); \hat{\boldsymbol{\sigma}}\right) \\ \text{subject to:} \quad & \text{crystallizer model} \end{aligned} \tag{NLP1}$$

The sequential quadratic programming code NPSOL [55] was used to solve this unconstrained optimization problem.

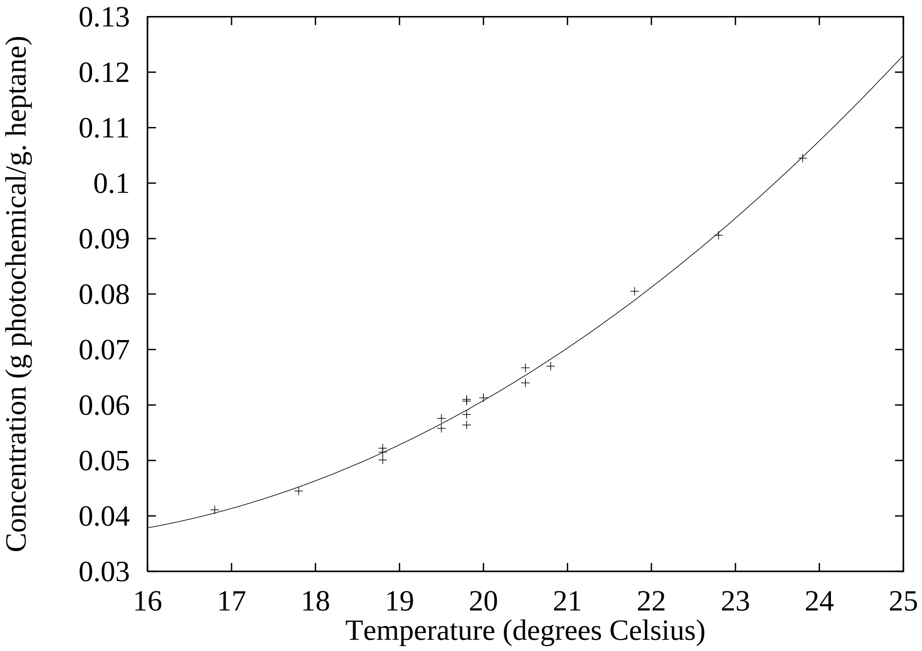

Figure 4.4: Solubility curve for the photochemical-heptane system.

4.4 Design of Initial Experiments

The designs of the initial batch crystallizer experiments were motivated both by the need to examine a reasonable range of operating conditions and to adhere to the limitations imposed by the model, the apparatus, and the chemical system. The key issues in the designs involved selection of the temperature range, the seed loads, and the cooling profiles.

The temperature range was selected to meet two conflicting goals; the first requiring that the span of the range be minimized to preserve the model assumption of temperature independent parameters and the second requiring that the range of crystallization driving force be maximized to increase the information content of the data. These issues were addressed by examining the solubility curve for the photochemical-heptane system, shown in Figure 4.4. The concentrations in Figure 4.4 were determined from calibrated densitometer data for crystal-solution equilibration times in excess of 24 hours. The curve is determined by fitting the data with a quadratic polynomial in temperature. The constants defining the solubility polynomial are given in Appendix A.

The sensitivity of the solubility curve decreases with temperature, suggesting that higher temperatures provide a wider range of driving force given a constant temperature span. However, because the amount of solute required to charge the crystallizer also increases with temperature, a range of $19 - 22°C$ was used to conserve the limited supply of raw photochemical. The initial concentration was chosen to be 0.11 grams of photochemical per gram of heptane, providing a $2°C$ cushion between the highest operating temperatures and saturation temperature of the initial solution. The combination of temperature range and initial concentration dictated initial relative supersaturations in excess of about 0.35. For most high purity, inorganic systems this high level of driving force would be well outside the metastable zone and could produce in spontaneous showers of nuclei. However, the photochemical-heptane system encounters no primary nucleation even at relative supersaturations above 1.0. The selected temperature range and initial concentration permits a maximum relative supersaturation of 1.2.

The seed masses for the initial runs were chosen to allow the collection of meaningful transmittance data. The crystallizer model uses the Beer-Lambert law to predict transmittance from the second moment of the CSD. It was noted in Section 2.5.3 that the applicability of Beer-Lambert is limited to transmittance values above about 0.8. Due to the small size and large surface-to-volume ratio of the photochemical seeds, seed masses above about 0.5 grams produce transmittance readings that violate the 0.8 limit. Therefore, the seed loads in the initial runs were all 0.15 grams or smaller to allow the second moment dynamics to be observed at early times in the run. However, within this limited range, the loads were varied and the effects of identical temperature profiles applied to different loads were examined.

The initial temperature profiles were chosen to represent traditional isothermal and linear cooling policies. These policies provided a quantitative evaluation of the crystallization kinetics over a wide range of driving forces. The final-time temperatures were varied to check the accuracy of the solubility. More exotic profiles that capitalize on specific dynamics of the system were then designed to acquire more informative data. The 4 profiles used in the initial set of experiments are given in Figure 4.5. The alphabetic listing corresponds to the order in which the profiles were implemented. Profile D was designed according to a D-optimal strategy (see Section 4.8.1) based on parameters estimated using data from replicate experiments with profiles A, B, and C. It is included in the initial set of experiments because it shares a data collection error with the initial set of runs. This error and the strategy used to correct it are explained in Section 4.5.3.

Figure 4.5 shows the actual crystallizer temperature data. The set point trajectories are not shown because they would be obscured completely by the data. The set point tracking is more than adequate for the purposes of this study.

4.5 Initial Parameter Estimates

The initial parameter estimates were determined from a set of nine batch crystallization experiments. As previously noted, the initial concentration in each run was 0.11 grams photochemical/gram heptane. The total mass of solvent was also constant with a value of

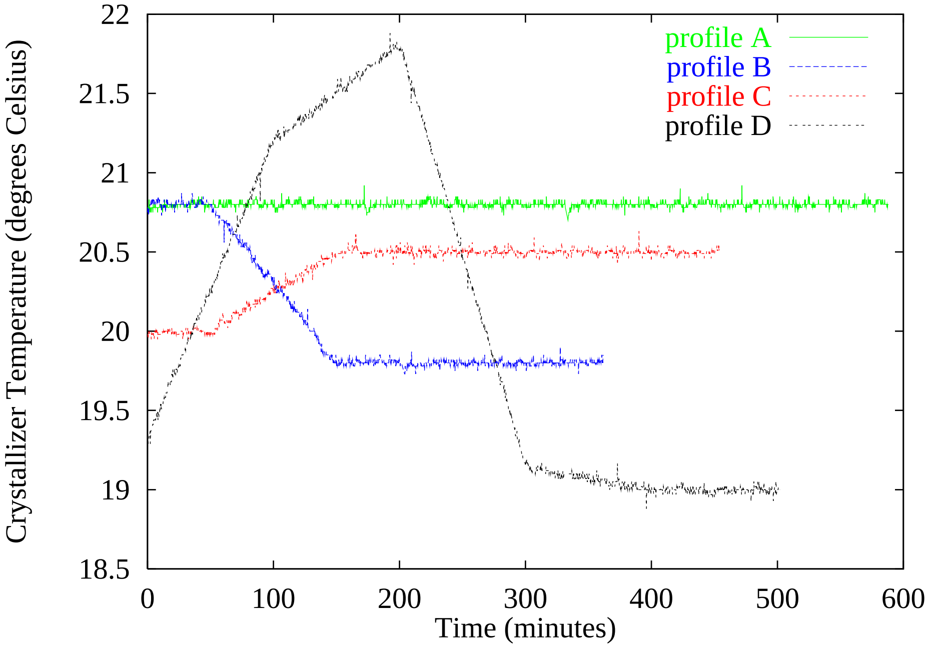

Figure 4.5: Representative temperature profiles from the initial set of nine batch crystallization experiments.

run number	seed mass	cooling policy
1a	0.0510 grams	profile A
1b	0.0518 grams	profile A
2a	0.1023 grams	profile B
2b	0.0983 grams	profile B
3a	0.1489 grams	profile C
3b	0.1527 grams	profile C
4a	0.0740 grams	profile C
4b	0.0745 grams	profile C
5a	0.1496 grams	profile D

Table 4.1: Seed loads and cooling profiles for the initial experiments.

1400.0 grams. The experiments were distinguished by their seed loads and cooling policies. These are given in in Table 4.1. The convention of differentiating experiments based on their seed loads and cooling policies was retained throughout the study.

The run naming convention used in Table 4.1 implies that runs 1a and 1b, 2a and 2b, and so on are replicates. Experimental replication is necessary to establish an upper bound on the expected quality of the model predictions. Of the nine initial experiments, there were five independent experiments, four of which were replicated. The only experimental differences between the replicates involve the random noise in the temperature profiles and the small seed load mismatches shown in Table 4.1.

Concentration and transmittance data from these nine runs were used to estimate the kinetic parameters according to NLP1. The optimal estimates selecting the mode of the posterior distribution are given, with their corresponding linear 95% confidence intervals, in Table 4.2. These estimates assume that the size distribution of the seeds is linear with an initial maximum size of 20 microns. The data and model fits are presented in Figures 4.6 through 4.11.

4.5.1 Comments on the Data

Before focusing on the quality of the model predictions, it is instructive to examine some features of the data. Figures 4.6 through 4.9 demonstrate the replicability of the experimental system. There was some mismatch between the concentration replicates of runs 1 and 2 but, in general, the replicate data sets were very similar. The model predictions of the replicates show some mismatch. This is especially clear in the run 1 transmittance

	$\ln(k_g)$	g	$\ln(k_b)$	b	j
Estimate	-1.08	2.05	25.24	2.10	2.26
Interval	± 0.03	± 0.04	± 0.35	± 0.25	± 0.08

Table 4.2: Parameter estimates and 95% confidence intervals for concentration and transmittance data from nine initial experiments.

predictions of Figure 4.6. For true replicates the model fits would cover each other. The disparity is indicative of the seed load and temperature profile noise noted earlier.

The data in Figures 4.8 and 4.9 represent runs 3 and 4. As indicated in Table 4.1, these runs were identical except for their seed loads. Run 3 had twice the seed mass of run 4. This doubling is clear in the initial transmittance measurements of Figure 4.8, where the run 3 replicates obscured approximately twice much light as did those of run 4 ($(I/I_0)_{\text{run3}} \approx 0.92$ versus $(I/I_0)_{\text{run4}} \approx 0.96$). Having twice the crystalline surface area available for deposition, the transfer of solute mass into the crystalline phase was accelerated for the run 3 replicates. This is shown by the faster decline in the concentration data for these runs.

For all of the data sets, the transmittance data were limited to a small period at the beginning of the crystallization. The transmittance data also exhibited more noise and had a lower sampling rate. These limitations resulted in parameter estimates that were determined largely through the reduction of errors in concentration.

4.5.2 Criticism of the Model Predictions

Examining the quality of the model predictions in Figures 4.6 through 4.11 reveals a clear pattern of small concentration residuals and very large transmittance residuals. This pattern implies that there is no set of kinetic parameters that allows the model structure to construct second and third moment trajectories corresponding to the data. As a result, the estimation method essentially ignores the transmittance data and locates parameters that produce third moment trajectories minimizing errors in the concentration data.

In the presence of prediction errors of this type, further criticism is without merit until the major deficiency has been repaired. In particular, the linear confidence intervals

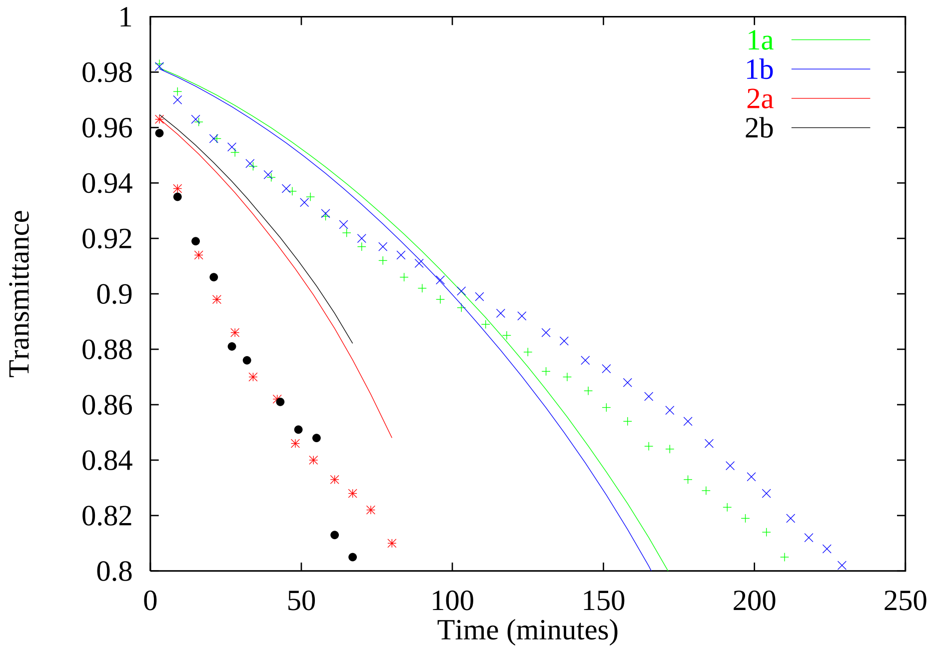

Figure 4.6: Transmittance profiles and model fits for runs 1a, 1b, 2a, and 2b, using the parameters in Table 4.2.

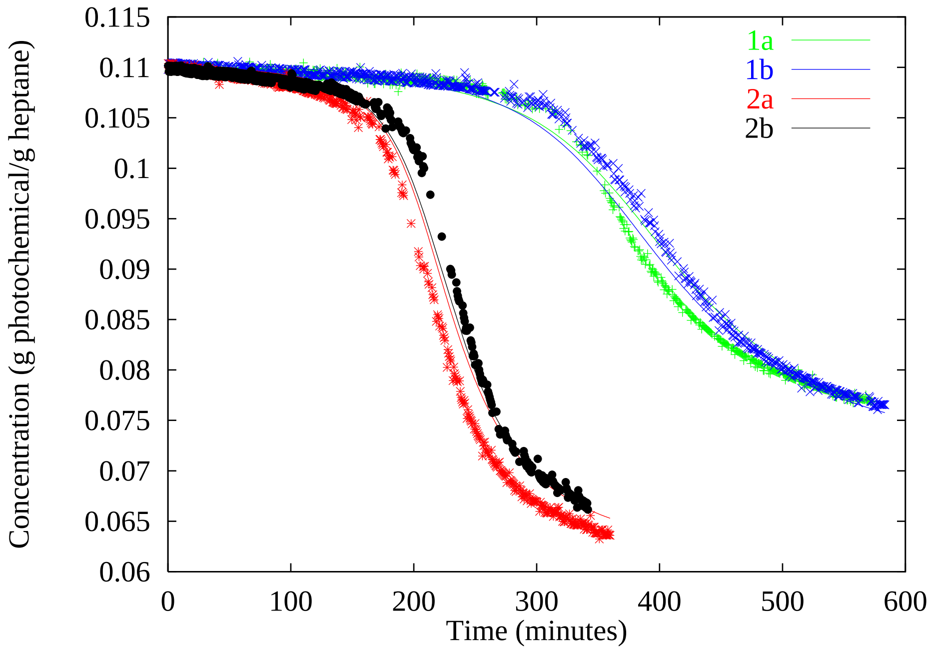

Figure 4.7: Concentration profiles and model fits for runs 1a, 1b, 2a, and 2b, using the parameters of 4.2.

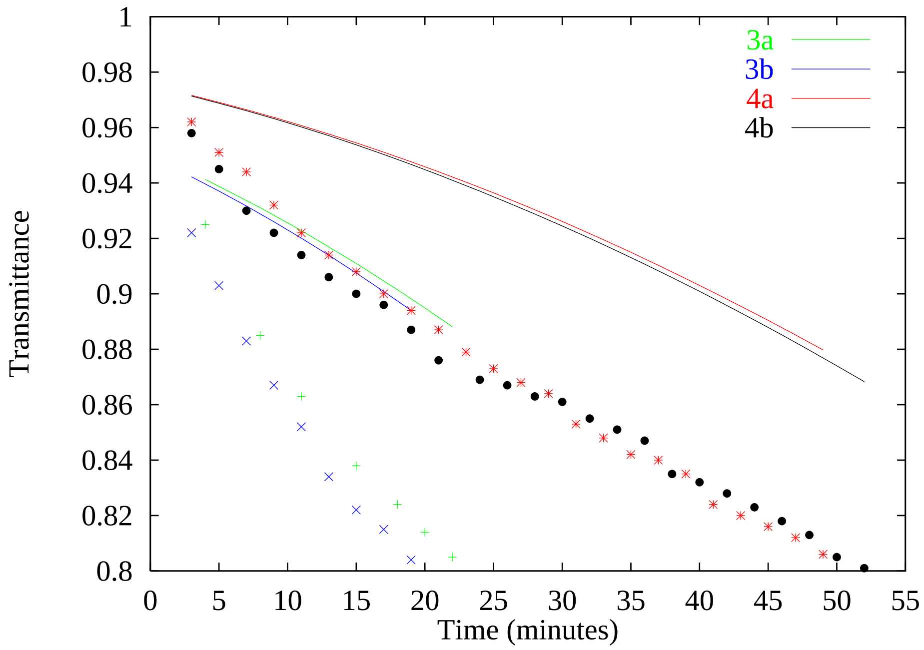

Figure 4.8: Transmittance profiles and model fits for runs 3a, 3b, 4a, and 4b, using the parameters in Table 4.2.

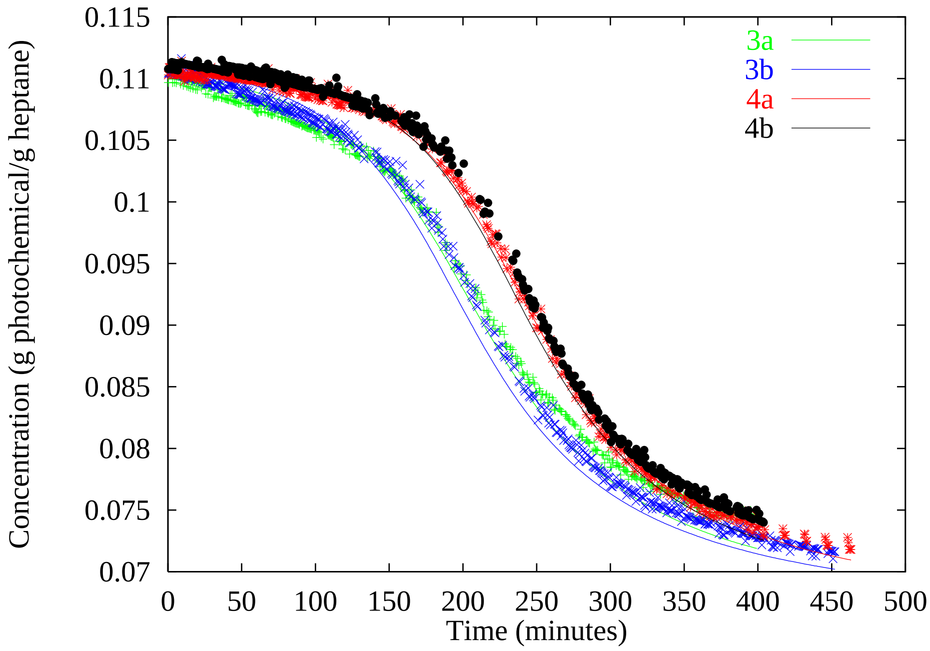

Figure 4.9: Concentration profiles and model fits for runs 3a, 3b, 4a, and 4b, using the parameters of 4.2.

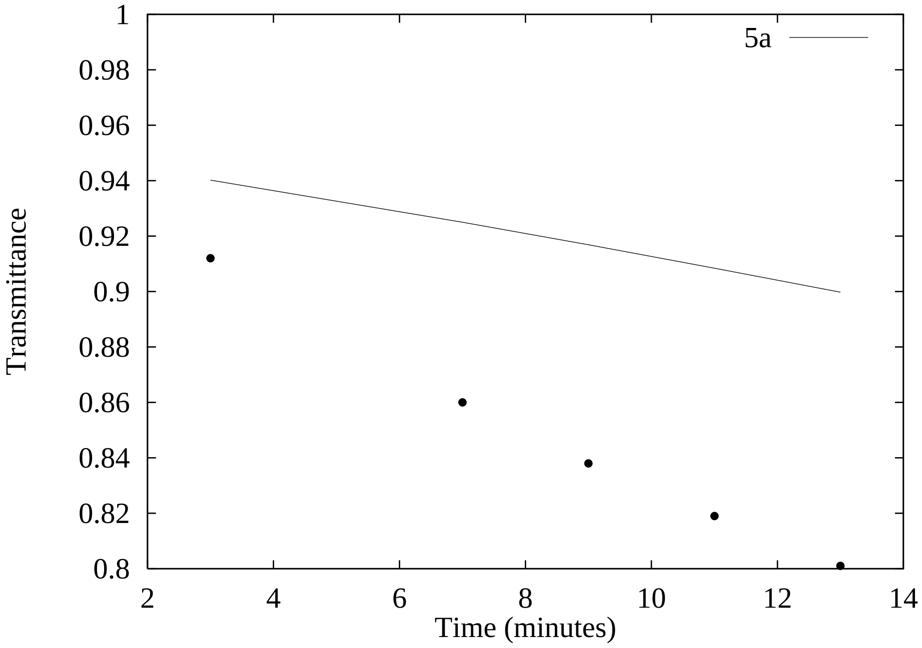

Figure 4.10: Transmittance profiles and model fits for run 5a, using the parameters in Table 4.2.

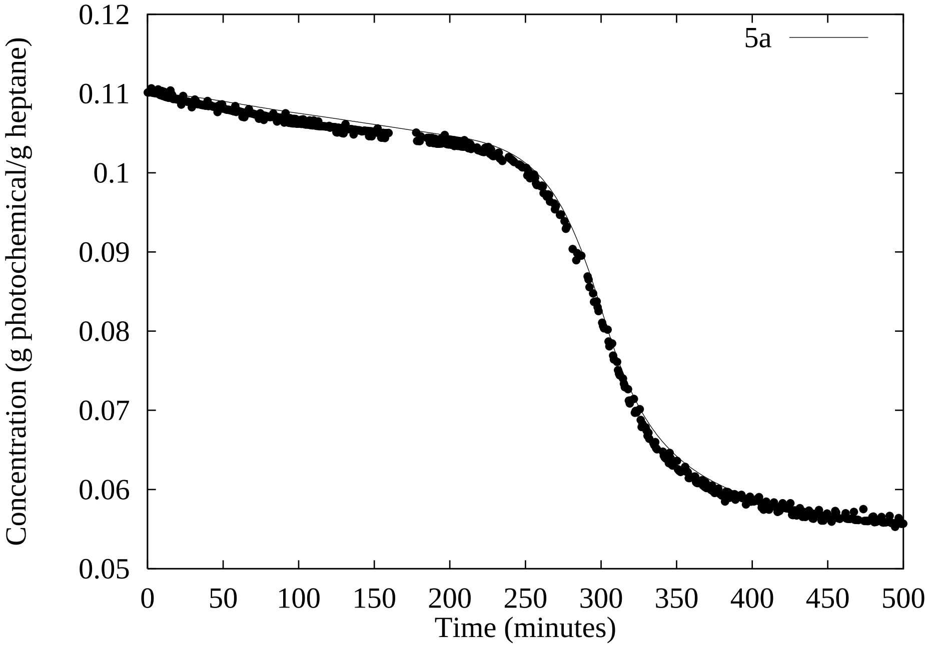

Figure 4.11: Concentration profiles and model fits for run 5a, using the parameters of 4.2.

presented in Table 4.2 are meaningless because the Gaussian Hessian approximation of Equation 4.23 is based on the assumption that the predictions errors are small.

4.5.3 Reassessing the Data

The inability to fit the transmittance data suggested that there were errors in the model, the data, or both. The first priority was to check the transmittance data. From a careful examination of the transmittance data, an anomaly was discovered that implied a possible bias. Note that with the exception of the run 1 replicates, the model always overpredicts the transmittance data. By contrast, for runs 1a and 1b the model underpredicts the data for a large portion of the run. It was proposed that this inconsistency could be due to the caking of crystals on face of the densitometer filter.

Caking problems of this type were discussed in Section 2.5.2. Essentially, the densitometer requires a solids-free sample stream to measure the concentration accurately. This is accomplished by using a mesh filter to keep crystals out of the stream. Over time, a cake of crystals may form on the face of the filter. Because the caked crystals are not mobile, they cannot access the transmittance sensor and a biased measurement results. Frequent back-flushing of the sample stream was used prevent the cake build-up in the initial experiments but this strategy may have been insufficient.

The run 1 replicates utilized the smallest seed loads and the lowest driving forces of all the runs. Thus, the evolution of the CSD was more gradual in these runs. There is reason to believe that a caking problem would have a more noticeable effect in such a situation because the disturbance could not be hidden by high nucleation and growth rates in the bulk slurry.

To test the bias theory, new replicates (i.e. runs with the same seed loads and temperature profiles) were performed for runs 1, 2, and 4. In these replicates, no concentration data were collected prior to the saturation of the transmittance sensor. Thus, there was no flow into the densitometer and the disturbance was removed entirely from the transmittance data. After the saturation of the transmittance sensor, concentration data was collected intermittently as described in Section 2.5.2. The seed loads and profiles of these new experiments are given in Table 4.3. Comparisons plots of all the run 1, 2, and 4 replicates are given in Figures 4.12 through 4.17.

run number	seed mass	cooling policy
1c	0.0517 grams	profile A
2c	0.1004 grams	profile B
4c	0.0731 grams	profile C

Table 4.3: Seed loads and cooling profiles for the transmittance bias experiments.

Clearly, the transmittance data collected for some of the initial experiments were erroneous. As shown in Figure 4.12, the most extreme errors occurred in the transmittance data for runs 1a and 1b. Caking on the filter resulted in a strong upward bias in the transmittance measurements. The caking also reduced the transfer of solute mass out of the solution phase. This concentration disturbance, shown in Figure 4.13, reflects the fact that the dissolved solute molecules did not have access to the surfaces of the caked crystals. These same effects are observable to a lesser degree in Figures 4.14 through 4.17 for runs 2 and 4.

As a result of these experiments, all of the data for runs 1a and 1b and the transmittance data from runs 2a and 2b were expunged from the estimation data set. These were

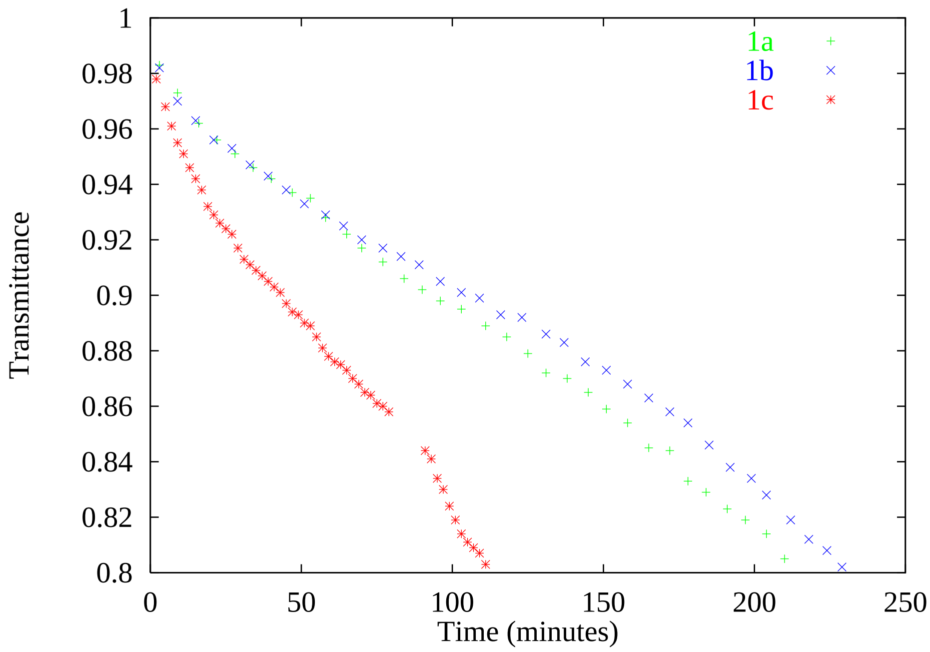

Figure 4.12: Transmittance profiles for runs 1a, 1b, and 1c.

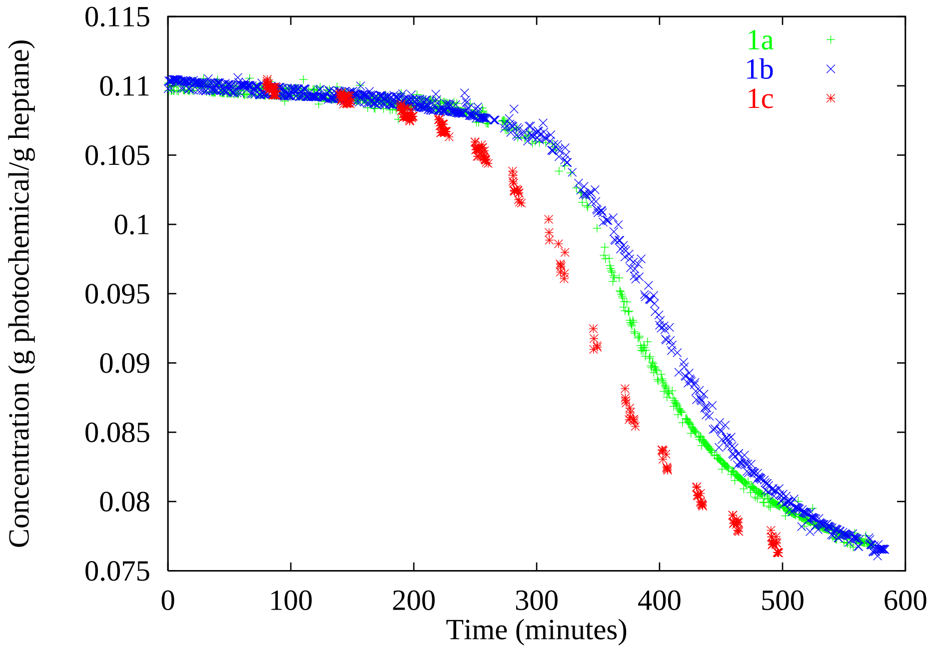

Figure 4.13: Concentration profiles runs 1a, 1b, and 1c.

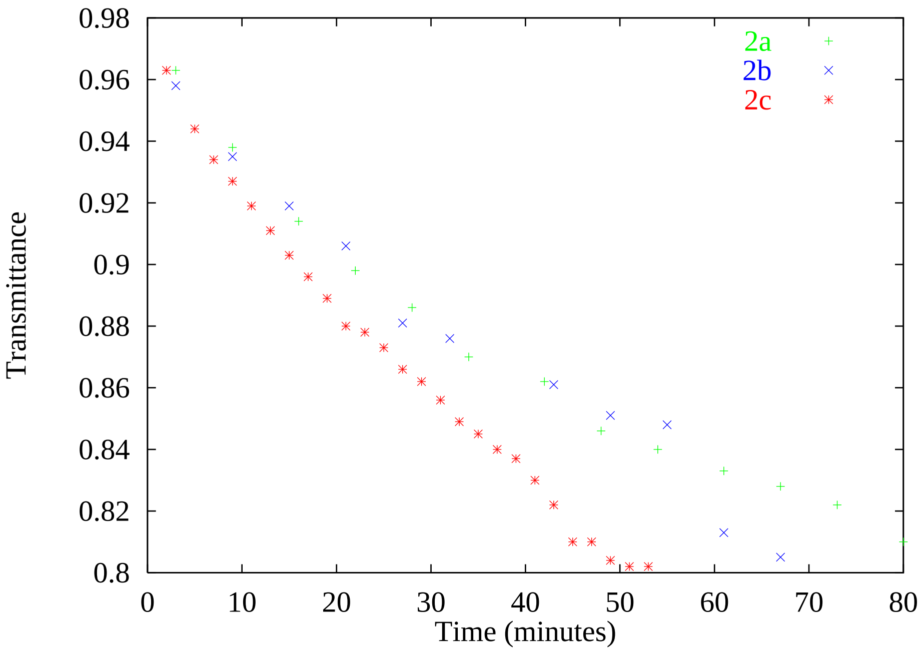

Figure 4.14: Transmittance profiles for runs 2a, 2b, and 2c.

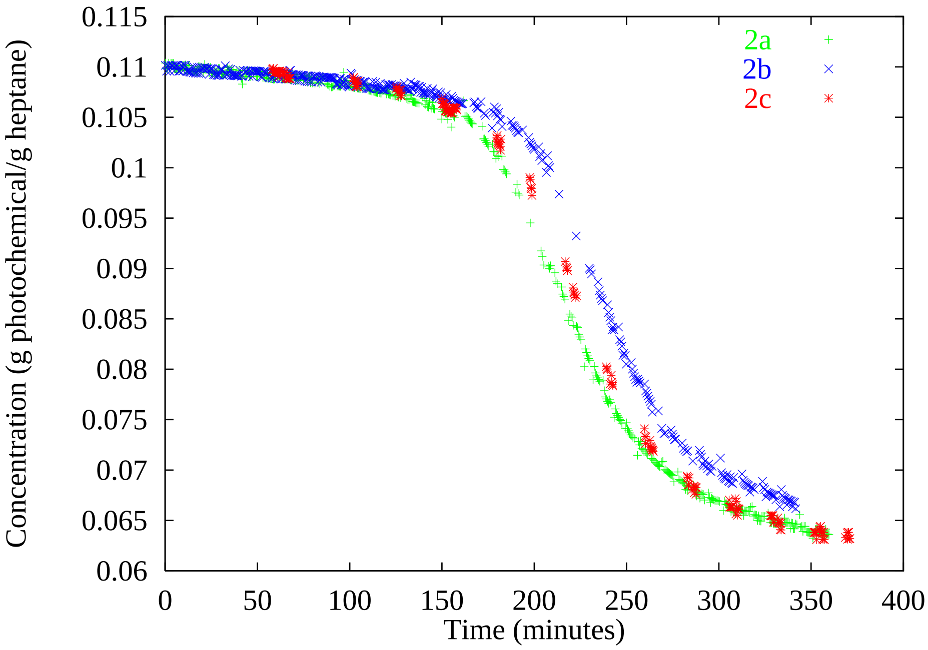

Figure 4.15: Concentration profiles runs 2a, 2b, and 2c.

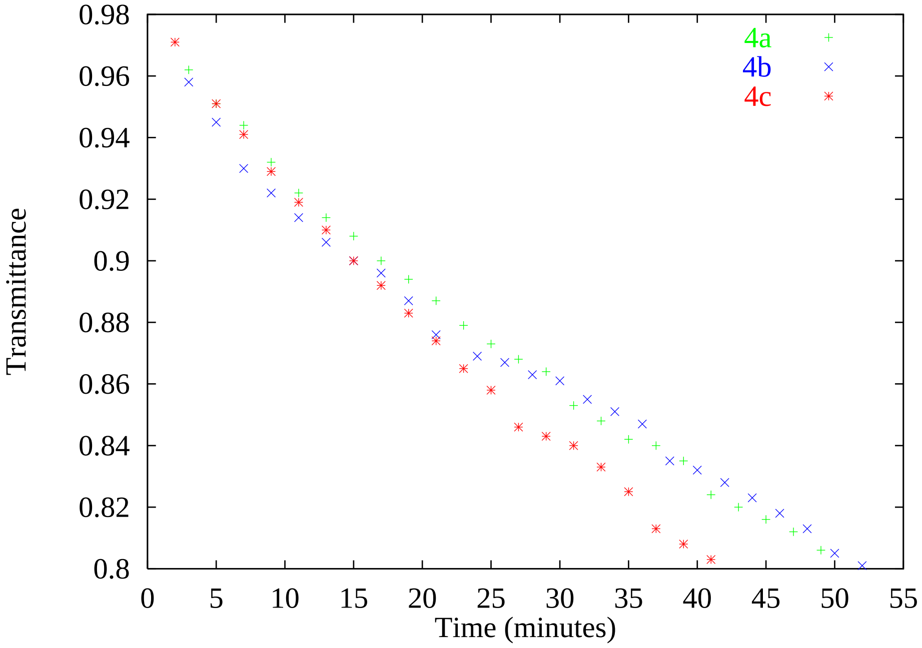

Figure 4.16: Transmittance profiles for runs 4a, 4b, and 4c.

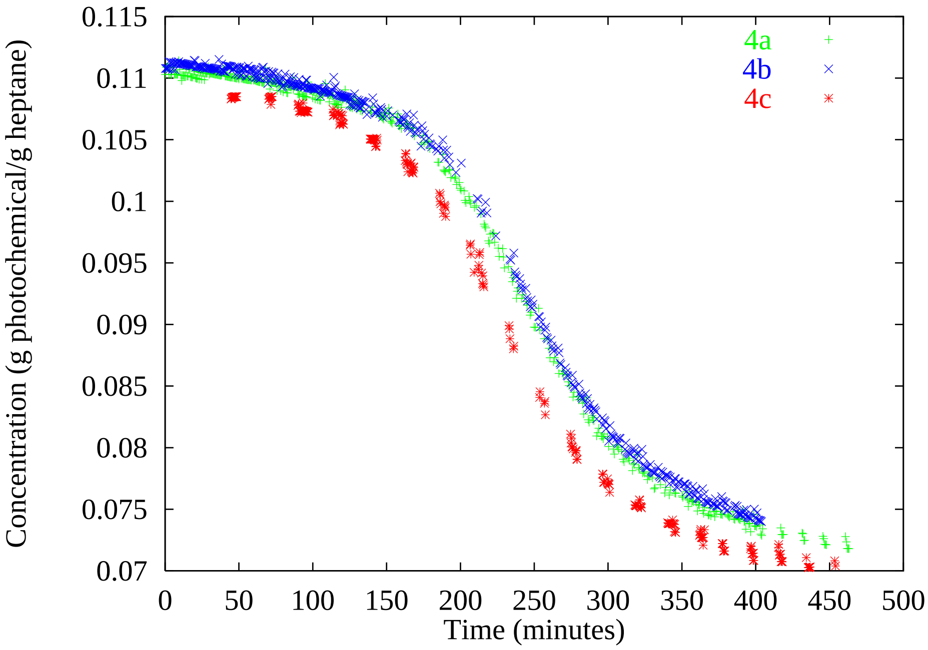

Figure 4.17: Concentration profiles runs 4a, 4b, and 4c.

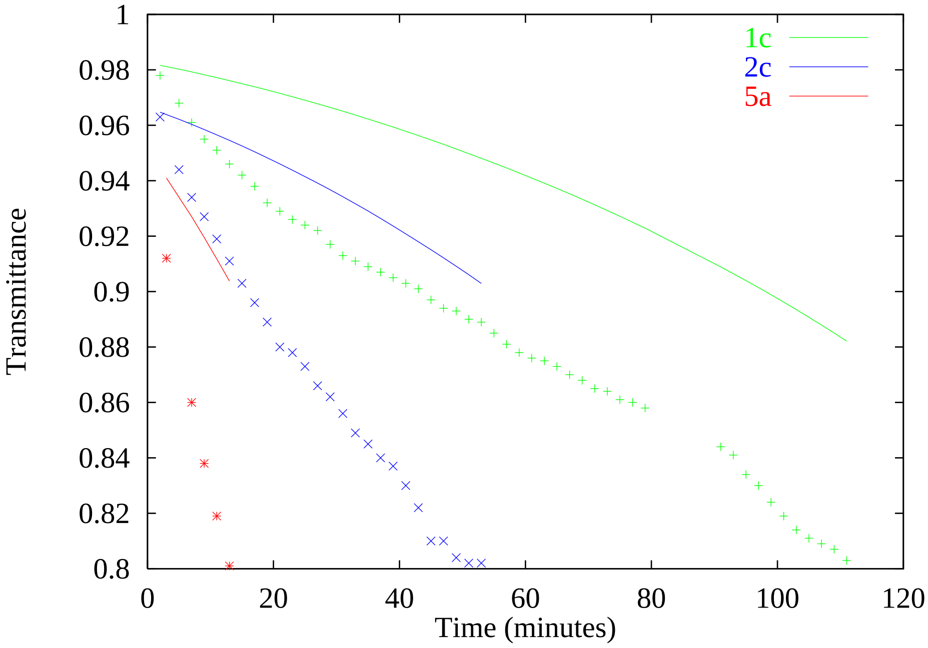

Figure 4.18: Transmittance profiles for runs 1c, 2c, and 5a.

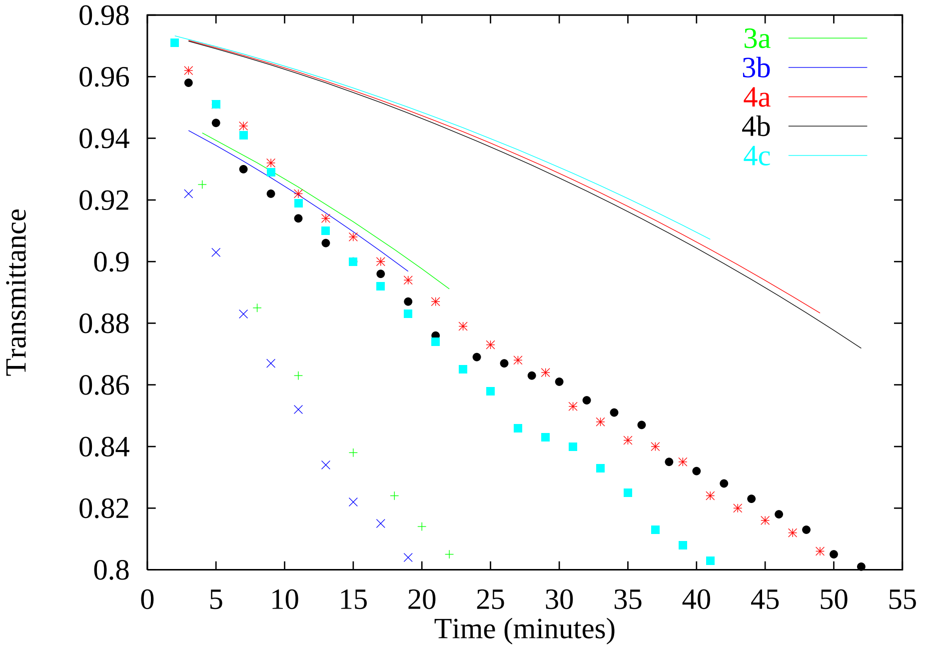

Figure 4.19: Transmittance profiles for runs 3a, 3b, 4a, 4b, and 4c.

replaced by the new data from runs 1c, 2c, and 4c. Using the revised data set, new parameters were estimated as before. However, the results remained unsatisfactory as, again, the concentration residuals were small and the transmittance residuals were large. The one improvement over the initial efforts was that the transmittance errors appeared to be consistent for all of the experiments. For the sake of brevity, only the transmittance fits are shown for this intermediate model (see Figures 4.18 and 4.19). Considering the large transmittance residuals there is, again, no reason to examine the estimates or confidence intervals in detail. Having already investigated the data, the revision of the model was considered next.

4.6 Model Revisions

The inadequacy of the model was characterized by large transmittance residuals and small concentration residuals. This incompatibility was addressed by examining the response models. These models are shown below,

$$\frac{d\hat{C}}{dt} = -3h\rho_c k_v G \mu_2 \tag{4.28}$$

$$\frac{I}{I_0} = \exp\left(-\frac{lk_a}{2}\mu_2\right) \tag{4.29}$$

where the ODE in Equation 4.28 is integrated numerically to provide predictions of $\hat{C}(t)$.

Under the initial assumptions, the only dynamic variables in Equations 4.28 and 4.29 are the growth rate and the second moment. Given the explicit role of the growth rate and the dependence of μ_2 on both growth and nucleation, the kinetic expressions were clear candidates for revision. However, another factor in both response models was the assumed habit of the crystals. Habit enters the concentration and transmittance models through the area and volume shape factors, k_a and k_v. These constants convert μ_2 and μ_3 into crystal surface area and volume, respectively. If the habit were characterized inaccurately or if there were unanticipated changes in habit during the initial crystallization runs, the faulty habit model could also be responsible for the prediction errors.

Two revisions were made to allow the general model to describe the crystallization data. A single parameter model was proposed to describe changes in the crystal habit of the seeds. In addition, the secondary nucleation model was extended to account for the sizes of the colliding crystals. The latter change was motivated by the fact that collisions involving small crystals do not produce nuclei because their collision energies are insufficient to dislodge structures from the crystal surfaces. Both revisions were required to attain adequate predictions of the data.

4.6.1 Modeling Crystal Habit

In the initial modeling effort, the habit of the seeds was assumed to remain constant and the nucleated crystals were assumed to have the same habit as the seeds. However, because the seed crystals may have been damaged during preparation (see Appendix A and Figure A.1), there was reason to believe that the average habit estimated from scanning electron micrographs (SEM) of the seeds was not maintained during subsequent growth. If this were the case, the initial habit would also be inadequate to describe nucleated crystals that were not affected by the seed preparations.

Neglecting breakage, crystal habit is determined by the relative growth rates of the individual faces of a crystal [97]. These growth rates are commensurate with how easily new solute molecules are incorporated into the crystal lattice structure of a particular

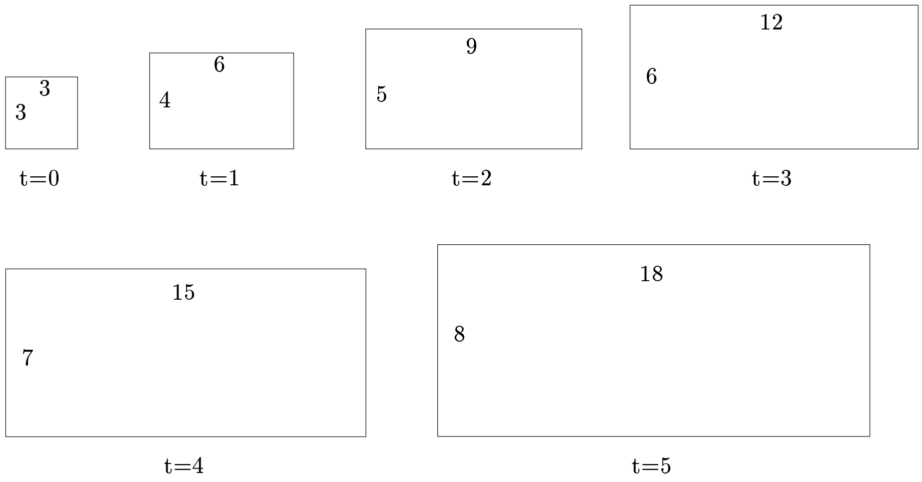

Figure 4.20: Two-dimensional representation of the growth of crystal faces. At each time step, the figure grows three units in length for every one unit in height. Note that the upper and lower *slow* faces become progressively longer than the left and right *fast* faces. As time increases, the figure approaches a length-to-height ratio asymptote of 3.

face. If one face grows quickly relative to others, the crystal dimension perpendicular to that face grows as well, resulting in larger facial areas for slower growing faces. This is illustrated in the simple two-dimensional example of Figure 4.20. If the shapes of the seed crystals were changed appreciably by the filtration or milling steps of their preparation, the initial estimates of k_a and k_v would not necessarily remain constant during subsequent crystallization because they would reflect the broken habit and not that corresponding to the relative growth rates of the faces. As a result, the habit could evolve from its initial state once the seeds were returned to a growth environment.

If breakage did occur during preparation, it most likely occurred perpendicular to the longest crystal dimension. Thus, the length of the seed crystals would be diminished but not depth or width. Therefore, if during subsequent crystallization the relative facial growth rates returned to the values that produced the seed crystals prior to breakage, a gradual increase in l_w (length-to-width ratio) would be expected. The changes in l_w would affect k_a and k_v according to Equations A.2 and A.3. Under the assumption that breakage events only diminished crystal length, the seed habit changes may be modeled as changes in l_w.

Given a seed with initial width, $L(0)$, a dynamic model of l_w for an individual crystal is

$$l_w(t) = \frac{l_w(0)L(0) + l_w(\infty)\Delta L(t)}{L(0) + \Delta L(t)} \tag{4.30}$$

where $l_w(0)$ is the initial length-to-width ratio determined from SEM or optical micrographs, $\Delta L(t)$ is the integral of the growth from $t = 0$ to t, and $l_w(\infty)$ is the assymptotic length-to-width ratio as ΔL goes to infinity. $l_w(\infty)$ is an estimate of the ratio of the growth rates along the length and width dimensions.

According to Equation 4.30, $l_w(t)$ is a function of the initial size of the seed crystal. An illustration of the effect of seed size on the evolution of $l_w(t)$ is given in Figure 4.21. This dependence may be integrated out by averaging Equation 4.30 over the initial size distribution of the seeds. The averaged value, $\overline{l_w}(t)$, is then a function of the integrated state, ΔL, and the constant parameters $l_w(0)$ and $l_w(\infty)$. The closed form of $\overline{l_w}(t)$ for a linear seed distribution of the type discussed in Section 3.7.1 is

$$\overline{l_w}(t) = l_w(0) + \frac{2\Delta L}{L_m}\left(l_w(0) - l_w(\infty)\right)\left[1 + \left(1 + \frac{\Delta L}{L_m}\right)\ln\left(\frac{\Delta L}{L_m + \Delta L}\right)\right] \tag{4.31}$$

where L_m is upper size limit of the linear seed distribution. The $\overline{l_w}(t)$ profile corresponding to Equation 4.31 with $L_m = 20$ microns is also presented in Figure 4.21.

The habit change model is implemented by computing ΔL along with the other states of the general model. Given a value $l_w(0)$, the instantaneous value of ΔL and the parameter $l_w(\infty)$ allow subsequent determination of $\overline{l_w}(t)$. The $\overline{l_w}(t)$ is then used to determine $k_a(t)$

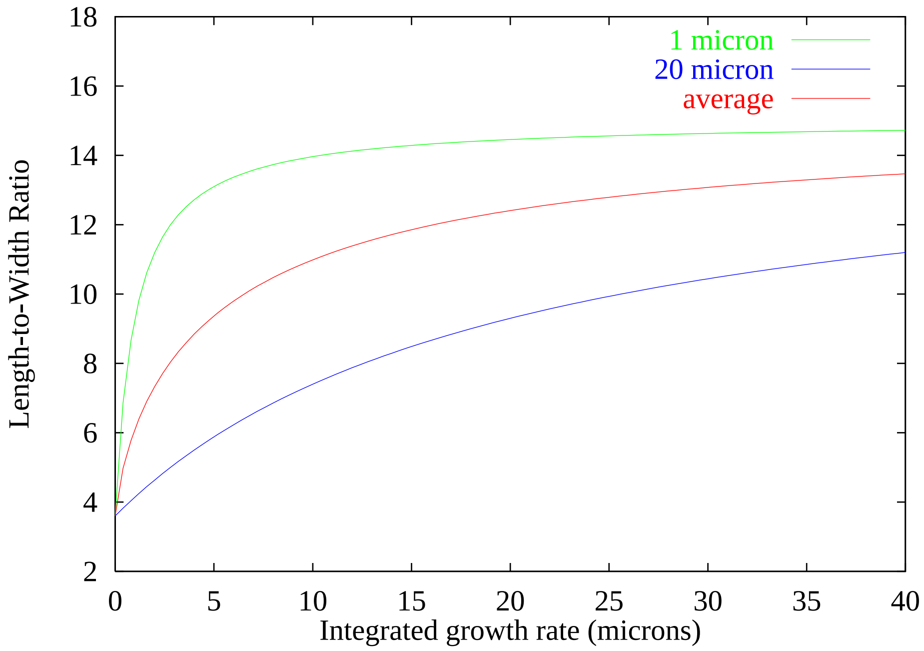

Figure 4.21: Dependence of length-to-width ratios on integrated growth rate. Evolution for a linear distribution averaged ratio is contrasted with profiles for individual crystals of 1 and 20 micron widths. The assymptotic limit as integrated growth rate goes to infinity is 15.

and $k_v(t)$ for the crystals in the seed mode via Equations A.2 and A.3. The nuclei were assumed to have the habit dictated by $l_w(\infty)$ as they were never damaged and were expected to develop habits corresponding to the facial growth rates. Since the shape factors affect directly the model predictions of the transmittance and concentration data and are functions of $l_w(\infty)$, $l_w(\infty)$ may be estimated from the data along with the kinetic parameters.

4.6.2 Size Dependent Nucleation

As discussed in Section 3.3, many different physical mechanisms have been proposed to describe secondary nucleation in solution crystallizers. The literature suggests that the key steps in the process are the breakage of structures on crystal surfaces due to crystal collisions and the subsequent regeneration of such structures via crystal growth. The initial nucleation model combined these two effects by relating nucleation to third moment of the CSD and to the relative supersaturation driving force.

$$B^{\circ} = k_b S^b \mu_3^j$$

The nucleation rate was correlated with the third moment to imply that the energy of collision is a crucial factor in the production of secondary nuclei. This model is closely related to those first proposed by Ottens et al. [104]. However, Ottens' models (see Equations 3.20a and 3.20b) also accounted for the fact that larger crystals are both more likely to collide and more likely to produce nuclei than small crystals due to their larger masses. As noted in Section 3.3, several researchers have found experimental evidence to support this hypothesis. This evidence suggested that the initial nucleation model may be improved by replacing μ_3 in the nucleation expression with a size-dependent value. The size-dependent third moment is denoted as μ_{3L} and is defined as the third moment of crystals larger than some minimum size $L_{\min}$.

In general, it is difficult to incorporate size dependence into a moment based model. However, because the shape of the seed size distribution is assumed to be known, it is possible to determine μ_{3L} for the seed class given the seed size distribution, the integrated growth rate ΔL, and the parameter $L_{\min}$. A similar development was not available for the nucleated crystals. Therefore, it was assumed that no nucleated crystals participated nuclei producing collisions (i.e. there were no second generation nuclei).

Figure 4.22 gives a schematic of how μ_{3L} is calculated for a linear seed distribution. As the crystallization progresses, the seed distribution is shifted ΔL units outward along the

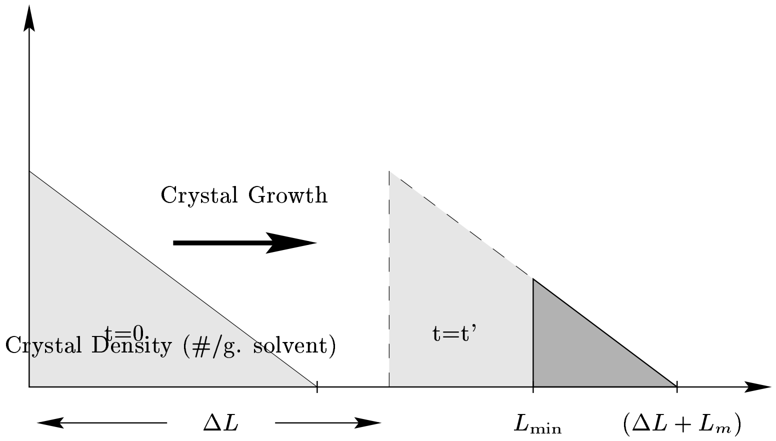

Figure 4.22: Growth of a linear seed size distribution through $L_{\min}$. The dark shading indicates the portion of the distribution included in the calculation of μ_{3L}.

size axis. The seed distribution may be rewritten to account for this shift by substituting $(L - \Delta L)$ for L in $\overline{f_s}(L)$ and by adding ΔL to the minimum and maximum sizes that specify the nonzero limits of the distributions. For example, the initial linear seed distribution has the equation

$$\overline{f_s(0)} = \begin{cases} f_l\left(1 - \frac{L}{L_m}\right), & \text{for } 0 \leq L \leq L_m \\ 0, & \text{otherwise} \end{cases}$$

which, in shifted form, becomes

$$\overline{f_s(t)} = \begin{cases} f_l\left(1 - \frac{L-\Delta L}{L_m}\right), & \text{for } \Delta L \leq L \leq (\Delta L + L_m) \\ 0, & \text{otherwise} \end{cases} .$$

μ_{3L} is then determined by multiplying the seed distribution by L^3 and integrating the product over the size range greater than $L_{\min}$. $L_{\min}$, in units of microns, is an adjustable parameter to be estimated from the transmittance and concentration data.

4.7 Parameter Estimates for Revised Model

The revised crystallizer model incorporated the dynamic habit and size-dependent nucleation submodels of Sections 4.6.1 and 4.6.2. The revised parameter vector contained seven parameters:

$$\boldsymbol{\theta} = \begin{bmatrix} k_g & : & \text{growth rate constant} \\ g & : & \text{growth rate order} \\ k_b & : & \text{nucleation rate constant} \\ b & : & \text{nucleation rate order} \\ j & : & \text{nucleation crystal mass order} \\ L_{\min} & : & \text{minimum crystal size for nucleation} \\ l_w(\infty) & : & \text{assymptotic length-to-width ratio} \end{bmatrix} \tag{4.32}$$

These parameters were estimated using transmittance and concentration data from ten crystallization experiments. The modal parameter estimates and their 95% linear confidence intervals are given in Table 4.4. The estimated prediction error variances are 2.31×10^{-6} for concentration and 2.66×10^{-3} for transmittance. The data sets and model predictions are given in Figures 4.23 through 4.28.

	$\ln(k_g)$	g	$\ln(k_b)$	b	j	$L_{\min}$	$l_w(\infty)$
Estimate	-1.50	1.89	29.57	3.72	2.21	28.13	17.37
Interval	± 0.01	± 0.01	± 0.88	± 0.29	± 0.14	± 0.69	± 0.20

Table 4.4: Parameter estimates and 95% confidence intervals for concentration and transmittance data using the revised model.

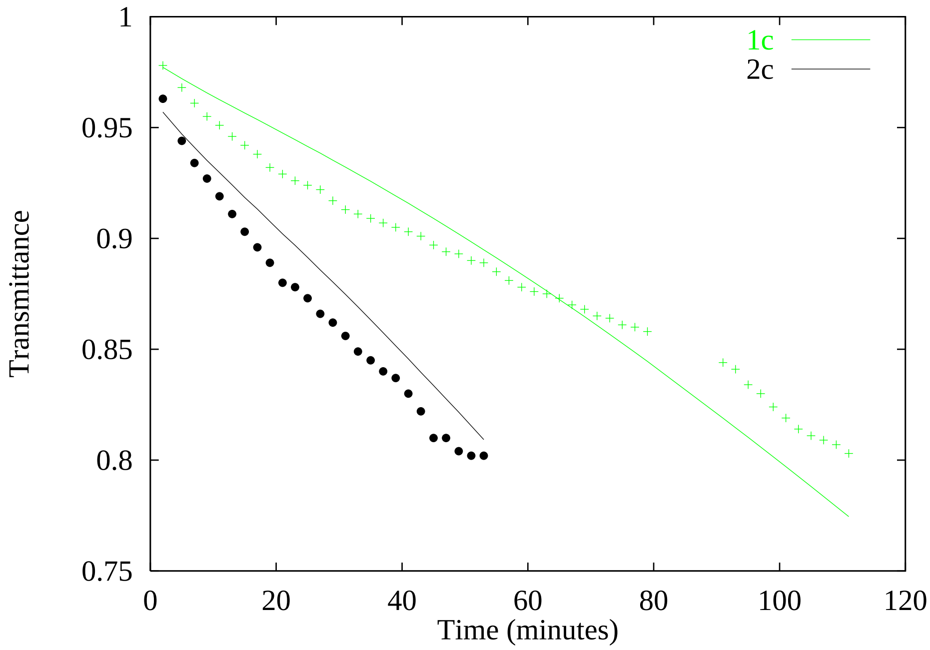

Figure 4.23: Transmittance profiles and model fits for runs 1c and 2c, using the parameters in Table 4.4.

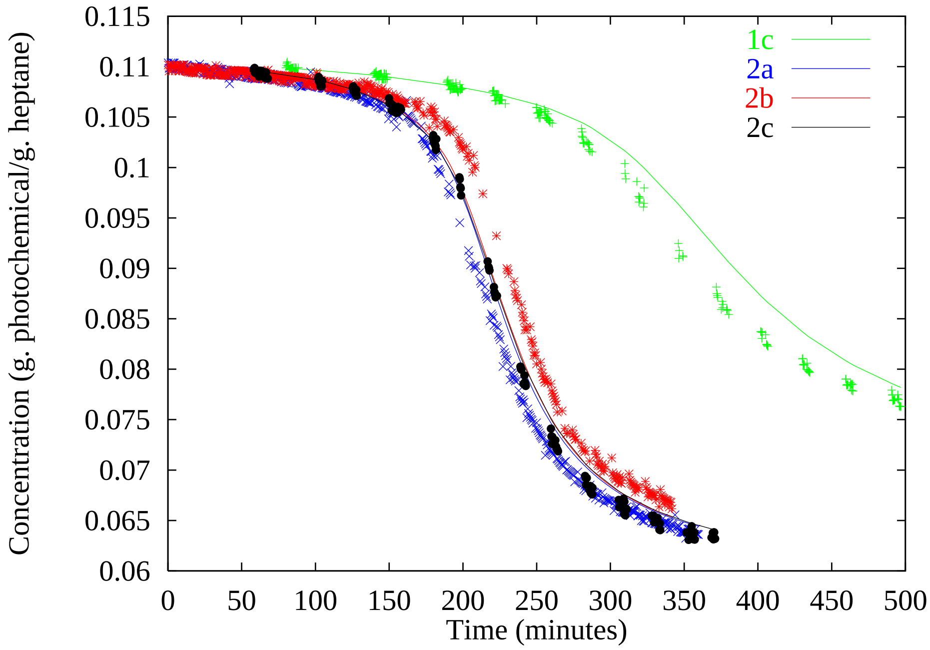

Figure 4.24: Concentration profiles and model fits for runs 1c, 2a, 2b, and 2c, using the parameters of 4.4.

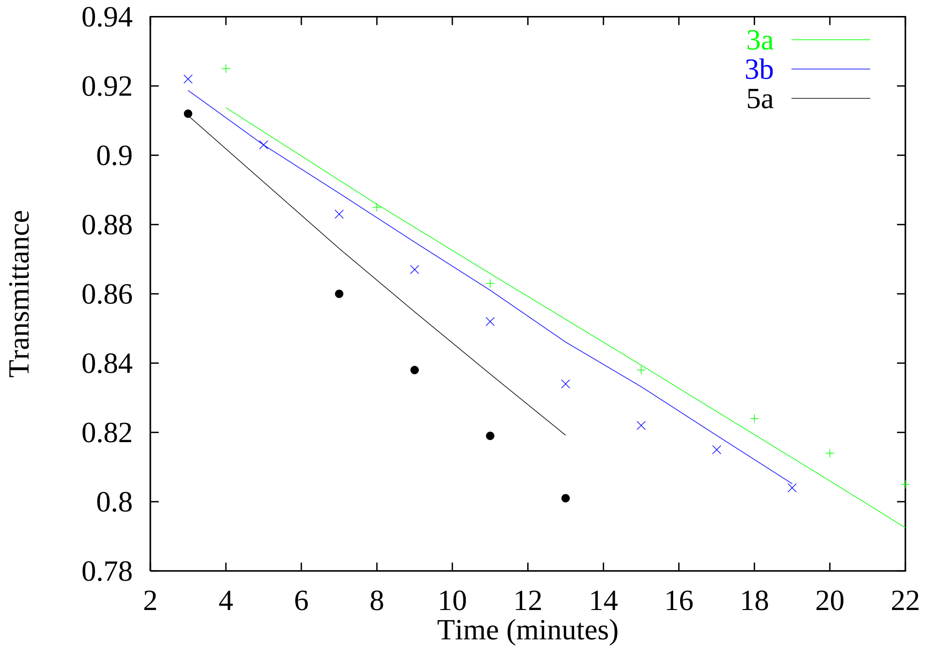

Figure 4.25: Transmittance profiles and model fits for runs 3a, 3b, and 5a, using the parameters in Table 4.4.

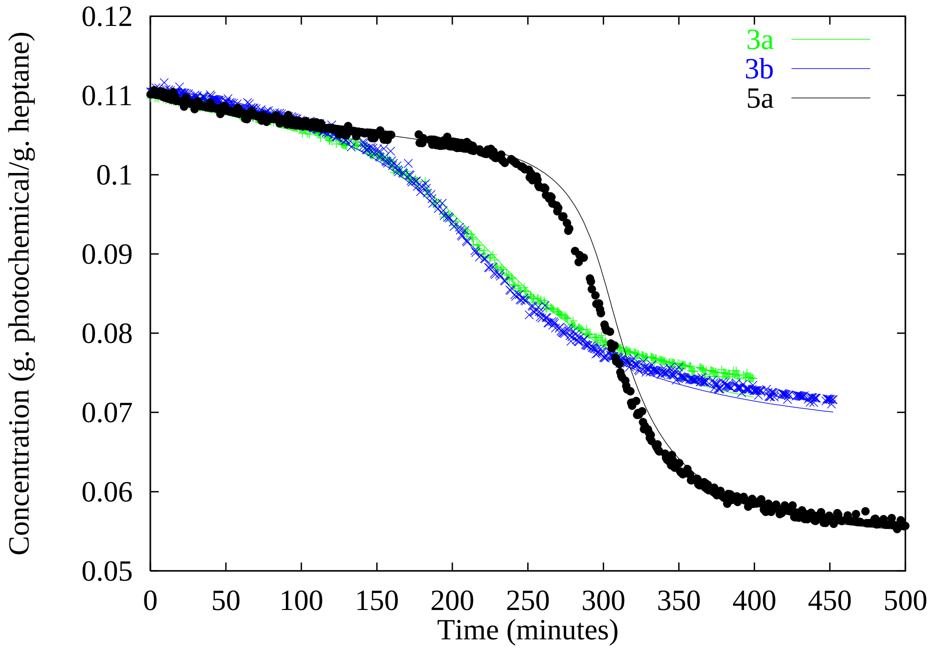

Figure 4.26: Concentration profiles and model fits for runs 3a, 3b, and 5a, using the parameters of 4.4.

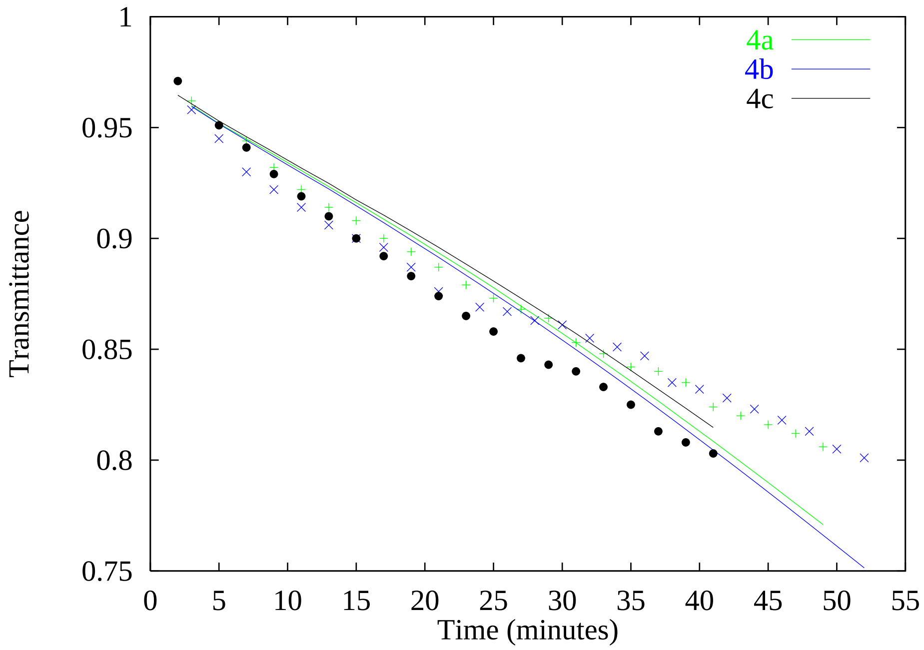

Figure 4.27: Transmittance profiles and model fits for runs 4a, 4b, and 4c, using the parameters in Table 4.4.

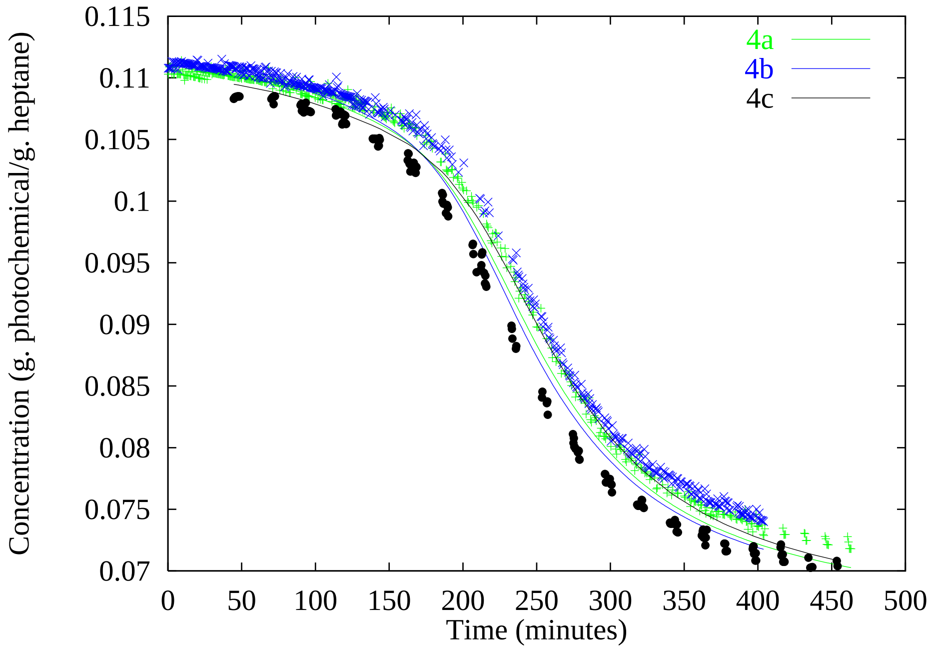

Figure 4.28: Concentration profiles and model fits for runs 4a, 4b, and 4c, using the parameters of 4.4.

4.7.1 Assessment of the Revised Model

From an examination of the model fits, the revisions clearly provide the model with a structure flexible enough to allow simultaneous prediction of both types of data. Given these improved fits, the linear confidence intervals in Table 4.4 are more likely to be meaningful than those determined prior to the revisions. The intervals indicate that the parameters are well-determined. The least certain parameters are b and j with intervals representing 7.8 and 6.3% of the parameter magnitudes, respectively. The growth parameters are the most precise with relative intervals of less than 1%.

By comparing the optimal parameter estimates with proposed mechanisms of growth and nucleation, it is possible to gain insight into the crystallization process. The growth order of relative supersaturation, g, is 1.90. This is close to the upper limit of 2 predicted by the BCF model (see Section 3.3) and implies that crystallization is limited by the integration of solute molecules into the crystal lattice and not by diffusion from the bulk. The supersaturation order of growth, g, is also significantly lower than the corresponding supersaturation order of the nucleation rate, b. That b is larger than g is a pivotal point with regard to optimizing the performance of the system as b and g must be appreciably different for relative supersaturation to serve as an effective variable to control the final-time CSD. That g is less than b indicates that high levels of supersaturation favor the nucleation process relative to that of growth. Since large nuclei populations hamper solid-liquid separation steps, the lower growth order suggests operating policies that minimize supersaturation levels facilitate filtration processes.

The third moment order of nucleation, j, is 2.20, indicating that crystal-crystal collisions are more influential than crystal-agitator contacts in spawning nuclei [149]. When crystal-agitator contacts dominate, j is found to be closer to one than to two. Values near 2 are rare but have been observed in lab-scale crystallizers [149].

On the basis of the improved fit, the revised crystallizer model was taken to be the correct structural form of the model. This assumption is supported by additional empirical evidence to be presented in a Section 4.9. Fixing the form of the model allowed the design of the final identification experiment to focus on the minimization of parameter uncertainty.

4.8 Optimal Experimental Design

As noted in Section 4.2.2, experimental design techniques use an existing model to optimize future experimental conditions in some respect. The optimization usually involves either the minimization of parameter uncertainty in a chosen model or the selection of a "best" model from a group of rivals. These types of designs are termed *extremal design* and *model discrimination*, respectively. Having selected a model form, this study focuses exclusively on extremal designs.

The reader interested in discrimination designs is encouraged to consult the texts of Fedorov [41] and Atkinson [3] and the papers of Hunter and Reiner [72] and Box and Hill [13]. In a related article, Stewart et al. [142] discuss a Bayesian approach to model discrimination using existing data.

The goal of the extremal form of optimal experimental design is to determine, given the present model, the optimal operating conditions for the next experiment such that the data to be collected reduce parameter uncertainty to the greatest extent. Having determined the expected optimal conditions, the experiment is conducted and data collected. The new data set is added to the prior set, if one existed, and the parameters are reestimated from the enlarged set.

As with the parameter estimation formulation of Section 4.2, extremal experimental design may be presented within a Bayesian framework. As shown in Equation 4.6, Bayesian estimation involves the determination of the posterior probability density of the parameters given a prior and the data. The prior distribution is usually noninformative and the parameter inference problem is dominated by the information in the data. In experimental design, the goal remains the determination of the posterior density of the parameters, or some characteristic thereof, but instead of inferring parameters the operating conditions of the next experiment are inferred subject to an informative prior. The informative prior is the *a priori* probability density of the parameters and is usually a posterior distribution based on an initial set of experiments.

Assume that the prior distribution was constructed from n initial experiments, then the Bayesian framework for the experimental design of the $(n+1)$th experiment is

$$p_{n+1}(\hat{\boldsymbol{\psi}}|\boldsymbol{Y}_{n+1}) \propto \ell(\hat{\boldsymbol{\psi}}|\boldsymbol{Y})p_n(\hat{\boldsymbol{\psi}}|\boldsymbol{Y}_n) \tag{4.33}$$

where the use of $\hat{\boldsymbol{\psi}}$ indicates that the likelihood and posterior distributions are distributions about the modal parameters of $p_n(\hat{\boldsymbol{\psi}}|\boldsymbol{Y}_n)$. The solution to the design problem is the set of experimental conditions that produce the data matrix $\boldsymbol{Y}$ that diminishes the spread of $p_{n+1}(\hat{\boldsymbol{\psi}}|\boldsymbol{Y}_{n+1})$ about $\hat{\boldsymbol{\psi}}$ to the greatest degree. Since the spread of a p-dimensional probability distribution is characterized either by the Hessian, $\boldsymbol{H}_{\hat{\psi}}$, or its inverse, $\boldsymbol{V}_{\hat{\psi}}$, the criteria used to evaluate the spread are typically functions of one or the other of these matrices.

Box and Hunter [14] show that given Normally distributed prediction errors and the independence assumptions of this study, the posterior Hessian for the $(n+1)$th experiment, $\left(\boldsymbol{H}_{\hat{\theta}}\right)_{n+1}$, may be determined according to

$$\left(\boldsymbol{H}_{\hat{\theta}}\right)_{n+1} = \left(\boldsymbol{H}_{\hat{\theta}}\right)_{n} + \sum_{k=1}^{m} \hat{\sigma}_k^{-2} \sum_{i=1}^{n'_k} \boldsymbol{B}_{ik}^T \boldsymbol{B}_{ik} \tag{4.34}$$

where $\left(\boldsymbol{H}_{\hat{\theta}}\right)_{n}$ is the Hessian of $p_n(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\sigma}}|\boldsymbol{Y}_n)$, the $\hat{\sigma}_k$ are the prior modal estimates of the prediction error variances, n'_k is the number of measurements of the kth type generated by the $\hat{\boldsymbol{\theta}}$ model, and the $\boldsymbol{B}_{ik}$ are defined by Equation 4.22. The $\boldsymbol{B}_{ik}$ vectors may be determined for data matrices generated using the assumed model structure, the prior modal parameter estimates, and the independent variable profiles that define prospective experiments. Thus, through Equation 4.34, the expected reduction in the uncertainty of the prior parameters may be cast in terms of the operating conditions of *future* experiments.

Note that in nonlinear models the $\boldsymbol{B}_{ik}$ are inseparable functions of the parameters and the iterative process of estimation and optimal experimental design may yield different designs as the modal parameters change with the addition of new data. In linear systems, or in nonlinear systems in which the parameters have converged, subsequent designs are identical, leading to repeated collection of data within an informative but limited range of the operating region. This situation motivates Stewart et al. [141] to emphasize the importance of not beginning model-based design too early for fear of identifying model parameters that are only appropriate within the limited operating region. Design repetition may prevent the discovery of structural model errors that are significant within the desired operating region of the process but are not apparent within the subregion charted by the optimal design. A possible solution to this problem is to use different criteria over the course of an estimation study.

4.8.1 Extremal Design Criteria

Given a means of calculating $\left(\boldsymbol{H}_{\hat{\theta}}\right)_{n+1}$, an optimization criterion must be selected. The criterion may be formulated to apply to a subset of the parameters or may be a general measure of the uncertainty in the entire parameter set. However, since subset criteria may be derived by applying the general criteria to the appropriate submatrices of $\left(\boldsymbol{V}_{\hat{\theta}}\right)_{n+1}$ [41] only general measures are presented below. For convenience and consistency, the $(n+1)$ subscript is dropped and the criteria are discussed in terms of the posterior parameter covariance matrix $\boldsymbol{V}_{\hat{\theta}}$.

General criteria, dubbed "portmanteau criteria" by Box and Hunter [14], are lumped scalar measures of the elements of the posterior covariance matrix. Portmanteau criteria summarize uncertainty information regarding all of the estimated parameters. The best known of these criteria is *D-Optimal design.* D-Optimal designs minimize the volume of the elliptical HPD regions discussed in Section 4.2.3. The mathematical equivalent of volume minimization is the minimization of the determinant of $\boldsymbol{V}_{\hat{\theta}}$. This is seen by generalizing the two-dimensional HPD ellipse of Figure 4.29 to p dimensions and noting that

$$\text{ellipse volume} \propto \prod_{i=1}^{p} \sqrt{\chi_p^2(\alpha)\lambda_i} = \sqrt{\chi_p^2(\alpha)^p \left|\boldsymbol{V}_{\hat{\theta}}\right|}$$

where λ_i is the ith eigenvalue of $\boldsymbol{V}_{\hat{\theta}}$. Minimizing the linear HPD ellipse corresponds to a sharpening of the posterior distribution of the parameters about its mode. Wald [153] first suggested this criterion for use with linear models and it was later advocated for nonlinear

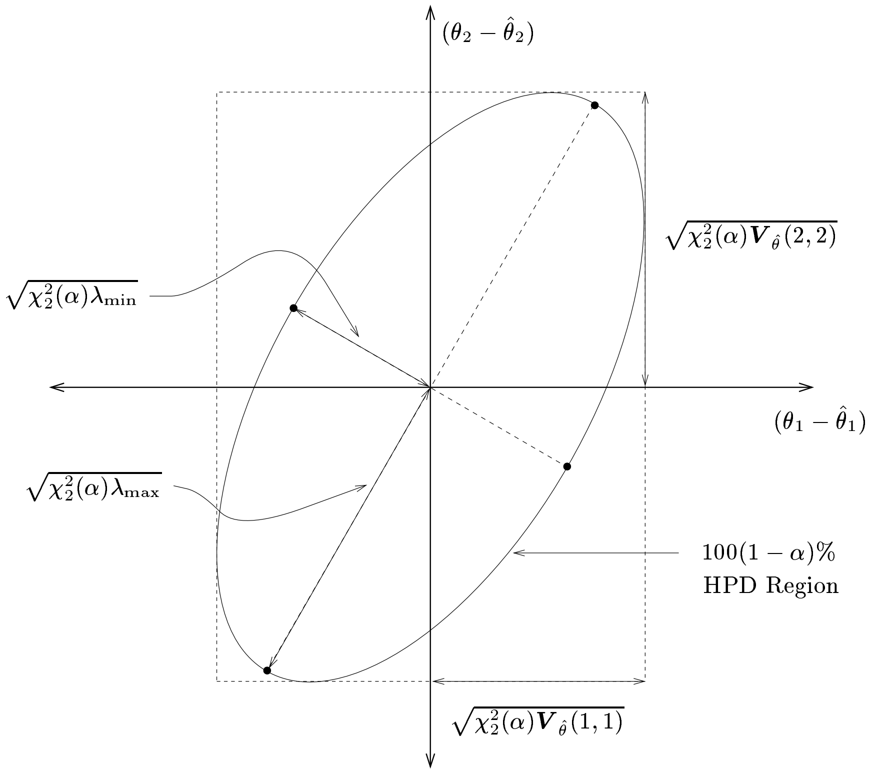

Figure 4.29: Elliptical region of highest posterior density (HPD) for a two parameter system. Conservative confidence intervals are defined by the circumscribing box. The lengths of the axes are proportional to the eigenvalues of $\boldsymbol{V}_{\hat{\theta}}$.

designs by Box and Lucas [12].

A second portmanteau criterion is *A-Optimal design*. A-Optimal designs minimize the arithmetic average of the parameter variances by minimizing the trace of $\boldsymbol{V}_{\hat{\theta}}$. Geometrically, the A-Optimal criterion minimizes the sum of the squared dimensions of the linear confidence interval hyperbox. The dimensions of the two-dimensional box are given in Figure 4.29. Box and Hunter [14] criticize the A-Optimal criterion because, unlike the determinant of D-Optimality, the trace of the posterior covariance is a function of the magnitudes of the individual parameters. Pinto et al. [109] repair this deficiency by using a revised trace criterion in which the ith diagonal element of $\boldsymbol{V}_{\hat{\theta}}$ is normalized by the square of the ith parameter.

Some researchers [65, 110] note that the D-Optimal measure does not account for the conditioning of the posterior covariance matrix. Hosten [65] observes that the volumes of HPD regions may be reduced significantly by experiments that shrink the shortest axis of the ellipsoid, thereby satisfying the D-Optimality criterion without improving the estimates of the least well-determined parameters. Hosten instead proposes the minimization of the largest eigenvalue of $\boldsymbol{V}_{\hat{\theta}}$ ($\lambda_{\max}$ in Figure 4.29). This criterion focuses the design exclusively on the reduction of uncertainty for the most poorly determined parameter and is, therefore, not a portmanteau criterion. Atkinson and Donev [3] refer to Hosten's criterion as the E-Optimal criterion.[3] Pritchard and Bacon [110] suggest augmenting D-Optimal design sequences with designs that minimize the sum of the squares of the off-diagonal terms in the parameter correlation matrix. The elements of this matrix are defined as

$$\frac{\boldsymbol{V}_{\hat{\theta}}(i,j)}{\sqrt{\boldsymbol{V}_{\hat{\theta}}(i,i)\boldsymbol{V}_{\hat{\theta}}(j,j)}} \tag{4.35}$$

where $\boldsymbol{V}_{\hat{\theta}}(i,j)$ is the ith row, jth column element of the parameter covariance matrix.

Of the criteria reviewed above, the general criteria of D-Optimal and normalized A-Optimal design are preferred because they account for the uncertainty of the entire set of estimated parameters [141].

4.8.2 Applications in Chemical Engineering

There have been many applications of extremal experimental design in chemical engineering fields but most have involved D-Optimal designs for sequential sets of steady-state

[3] The mnemonic devices for the alphabetic designs are "D for determinant", "A for average variance", and "E for extreme".

experiments. Examples of such efforts include those of Reilly et al. [128] and Sutton and MacGregor [143]. The former involves the application of design to improve estimates in first order kinetic models, while the latter examines the use of optimal design to improve the information content in vapor-liquid equilibrium experiments. Sutton and MacGregor [143] present an interesting comparison of D-Optimal and E-Optimal designs with a correlation minimization criteria (Turing's measure of conditioning) that maximizes the sphericity of the interval ellipses. The authors note that E-Optimality provides a compromise design between the the volume and sphericity methods.

As noted previously, the experiments in this study are distinguished by their seed masses and temperature profiles. Preliminary design simulations indicated monotonic reductions in uncertainty with increasing seed mass, the optimal design used in this study sets the seed mass to high end of the range allowed by the transmittance model and infers an optimal temperature profile. By using the temperature profile to characterize the optimal experiment, the design computations are clearly dynamic and require the integration of the crystallizer model ODE's.

Optimal design in systems described by nonlinear ODE's is investigated by several authors. Hosten and Emig [66] examine the design of optimal but constant initial and feed conditions. These authors use simulated data sets to assess D-Optimal and E-Optimal criteria for determining optimal initial conditions for a simple batch reactor and feed conditions for an adiabatic tubular reactor. As is common in experimental design, all optimal designs are found to lie on the boundaries of the specified operating regions. After a series of sequential designs, the E-Optimal criterion is said to provide better parameter estimates on the basis of its improvement in the estimates of the most poorly determined kinetic parameters.

Other studies concern time varying input signals. Murray and Reiff [100] use simulated experiments to investigate the estimation of thermal transport properties in an adiabatic packed bed. The method infers the optimal period of a square-wave inlet temperature policy using both D-Optimal and E-Optimal criteria. Parameter uncertainty reductions are achieved, however, there is no appreciable difference between the optimal input periods suggested by the two methods.

Unlike previous studies that used combinatorial evaluations of rough grids to determine the optimal conditions, Espie and Macchietto [38] formulate experimental design as an optimal control problem. Similar to the nonlinear optimization-based parameter estimation method of Section 4.3, the nonlinear model is integrated at each iteration of an NLP. The decision variables are the parameters defining the experimental operating policy. Espie and Macchietto demonstrate the flexibility of the method by determining designs for model discrimination and uncertainty reduction in a simulated fermentation reactor. The manipulated variables are the periods and magnitudes of piecewise constant time profiles for feed flow rate and concentration.

In this study, optimal design of the final estimation experiment was implemented as an optimal control problem. The implementation is similar to that developed by Espie and Macchietto [38], except that the time varying system temperature was parameterized as a piecewise linear profile.

4.8.3 Optimal Experimental Design Calculations

As discussed above, optimal experimental design involves the minimization of a measure of parameter uncertainty with respect to the experimental operating policy. The optimal design problem may be cast as an nonlinear, open-loop optimal control problem having the following general form:

$$\begin{aligned} \min_{\boldsymbol{u}(t)} \quad & \Phi(\boldsymbol{u}(t), \mathbf{x}(t_f); \hat{\boldsymbol{\theta}}) \\ \text{subject to:} \quad & \text{Crystallizer Model} \\ & \boldsymbol{h}\left(\boldsymbol{u}(t), \mathbf{x}(t)\right) = \mathbf{0} \\ & \boldsymbol{g}\left(\boldsymbol{u}(t), \mathbf{x}(t)\right) \geq \mathbf{0} \end{aligned}$$

where $\boldsymbol{u}(t)$ is a vector of manipulated variable inputs, $\boldsymbol{x}(t_f)$ are the states of the model evaluated at the final time t_f, and $\boldsymbol{h}\left(\boldsymbol{u}(t), \mathbf{x}(t)\right)$ and $\boldsymbol{g}\left(\boldsymbol{u}(t), \boldsymbol{x}(t)\right)$ are vectors of nonlinear equality and inequality constraints, respectively. Given a finite dimensional parameterization of $\boldsymbol{u}(t)$, the control problem becomes a formal nonlinear program that may be solved by standard algorithms that allow the incorporation of simple bounds and linear constraints on $\boldsymbol{u}(t)$ and final-time constraints on $\mathbf{x}(t_f)$.

Consider the initial conditions (e.g. initial concentration, seed mass, etc.) and duration of an optimal crystallization experiment to be fixed. The input vector, $\boldsymbol{u}(t)$, defining the experimental design is then fully specified by the vector that defines its temperature profile. Designate this vector $\boldsymbol{T}$. If the N_T elements of $\boldsymbol{T}$ are the vertices defining a piecewise linear temperature profile and the constraints consist of simple bounds on the elements of $\boldsymbol{T}$, linear constraints on the cooling rate, a nonlinear path constraint on relative supersaturation, and final-time constraints on the yield, then the optimal design problem may be restated as the following NLP:

$$\begin{aligned} \min_{\boldsymbol{T}} \quad & \Phi_d(\boldsymbol{T}, \mathbf{x}(t_f); \hat{\boldsymbol{\theta}}) \\ \text{subject to:} \quad & \text{Crystallizer Model} \\ & \boldsymbol{b}_l \leq \boldsymbol{A}\boldsymbol{T} \leq \boldsymbol{b}_u \\ & S(t) \leq S_{\max} \\ & \hat{C}(t_f) \leq \hat{C}_{\min} \end{aligned} \tag{NLP2}$$

where $\hat{C}(t_f)$ is the solution solute concentration at t_f and $\Phi_d(\boldsymbol{T}, \mathbf{x}(t_f); \hat{\boldsymbol{\theta}})$ may be any of the design criteria introduced in Section 4.8.1. $\boldsymbol{A}$ is a $(2N_T - 1) \times N_T$ constant matrix mapping the N_T-dimensional vector $\boldsymbol{T}$ into a $(2N_T - 1)$-dimensional vector containing the elements of $\boldsymbol{T}$ and the corresponding cooling rates they define. Appropriately, the first N_T components of $\boldsymbol{b}_l$ and $\boldsymbol{b}_u$ specify simple bounds on $\boldsymbol{T}$ while the remaining elements specify simple bounds on the cooling rates. The sequential quadratic programming code NPSOL [55] was used to solve NLP2. Note that, as with the parameter estimation procedure, the model-based objective function is evaluated explicitly at each iteration of the NLP by the differential-algebraic solver DDASSL [108].

The constraints imposed in NLP2 assure the investigator that the optimal design conditions reside in a realistic operating range. The simple bounds and cooling rate constraints prevent the identification of designs that require temperatures or cooling rates beyond the capacity of the experimental apparatus. These are linear constraints that may be specified within NPSOL. The final-time yield constraint ($\hat{C}(t_f) \leq \hat{C}_{\min}$) assures that the optimal design crystallizes a nontrivial amount of solute mass. In general, this nonlinear final-time constraint is not active for optimal design strategies but it is of more concern in the related determination of temperature profiles that optimize filtration performance.

As with linear constraints, nonlinear final-time constraints may be specified explicitly in NPSOL.

Constraining the maximum value of S allows the optimal design to be maintained within a desired driving force range. This type of nonlinear path constraint may not be explicitly specified in NPSOL. Instead, the path constraint is applied via the method of Sargent and Sullivan [135]. This requires the specification of an additional state equation that integrates the squares of the constraint violations over the course of the run. By requiring this state to be zero at t_f, the path constraint is converted into a nonlinear final-time constraint similar to minimum yield. Additional information regarding constraint handling in NLP2 are given by Miller and Rawlings [94].

Given the model parameters in Table 4.4, D-Optimal and normalized A-Optimal [109] designs were investigated in this study. As noted by Box and Lucas [12], multiple local optima exist for general design criteria. Therefore, both optimal designs described below constitute the best results of a series of optimizations started from several different initial temperature profiles.

4.8.4 D-Optimal and A-Optimal Temperature Profiles

Prior to determining the optimal temperature profiles, initial conditions and a general parameterization of the temperature profiles needed to be specified. As noted above, initial work indicated that higher seed masses would provide more informative experiments. Thus, the seed mass was set to 0.25 grams, the highest mass for which an appreciable amount of transmittance data could be collected. As in earlier experiments, the initial concentration was set to 0.11 grams photochemical/gram heptane. A piecewise linear batch cooling profile, represented by the vector of vertice temperatures, $\mathbf{T}(t)$, was used to optimize the design criteria. This parameterization allowed the specification of distinct upper and lower temperature bounds of 23° and 19° Celsius, respectively. The time discretization of the profiles located vertices at the 0, 100, 200, 300, 400, and 500 minute points in a 500 minute experiment. These choices are arbitrary but reflect a desire to give the optimizer the flexibility to manipulate temperature during any portion of the run without requiring the estimation of a large set of decision variables. The maximum cooling rate was set to 0.1°C per minute and the maximum allowable relative supersaturation was set to 1.0.

In order for the simulated Hessians to predict accurately the expected parameter uncertainties the types, numbers, and times of the simulated data should closely approximate those that will be collected in the actual experiment. The D and A-Optimal designs were determined using concentration and transmittance measurements predicted once per minute over the course of a 500 minute simulated experiment. Transmittance predictions with values lower than the Beer-Lambert law limit of 0.8 were excluded from the calculation. Thus, the simulated data set closely approximated that of the previous crystallization experiments described in this study. Ultimately, the accuracy of the prediction also depends upon the degree of similarity between the prior and posterior parameter sets.

The optimal temperature profiles determined for the D and A-Optimal criteria are given, along with the profiles from previous crystallizations, in Figure 4.30. The D-Optimal profile is similar to the profile used in run 5a. Recall that run 5a was a D-Optimal design corresponding to a set of model parameters determined prior to the discovery of the densitometer-induced bias in the transmittance data. The differences in the two D-Optimal designs are due to differences in seed mass and the changes in the structure and parame-

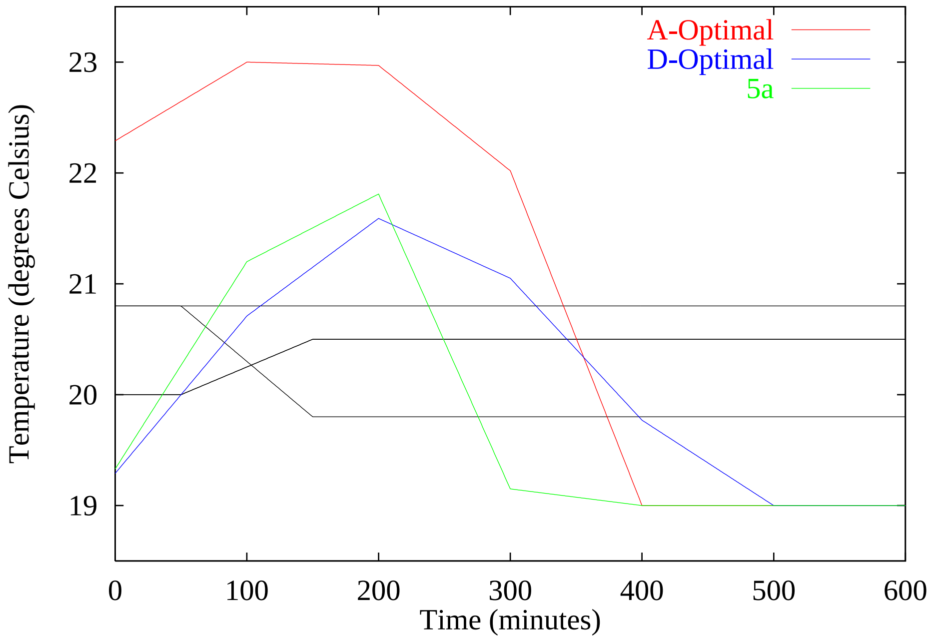

Figure 4.30: A-Optimal and D-Optimal temperature profiles as compared to initial profiles. Optimal profiles determined using the parameters in Table 4.4 with 0.25 gram seed mass.

ters of the crystallization model that followed from the removal of the biased data and the revisions of Sections 4.6.2 and 4.6.1.

The optimal experimental designs are defined in terms of their optimal temperature profiles. However, because the kinetic parameters correlate nucleation and growth to the relative supersaturation driving force, the supersaturation trajectories are the keys to understanding the optimal designs. The relative supersaturation trajectories, simulated from the parameters of Table 4.4, are given in Figure 4.31. Note that the relative supersaturation trajectories for both types of optimal profiles seek out regions of the operating space that are not covered by other experiments. Different operating regions indicate that the response models are most sensitive to the parameters in areas of the operating space that have not been charted by previous experiments. In order to maximize the data to be gathered in

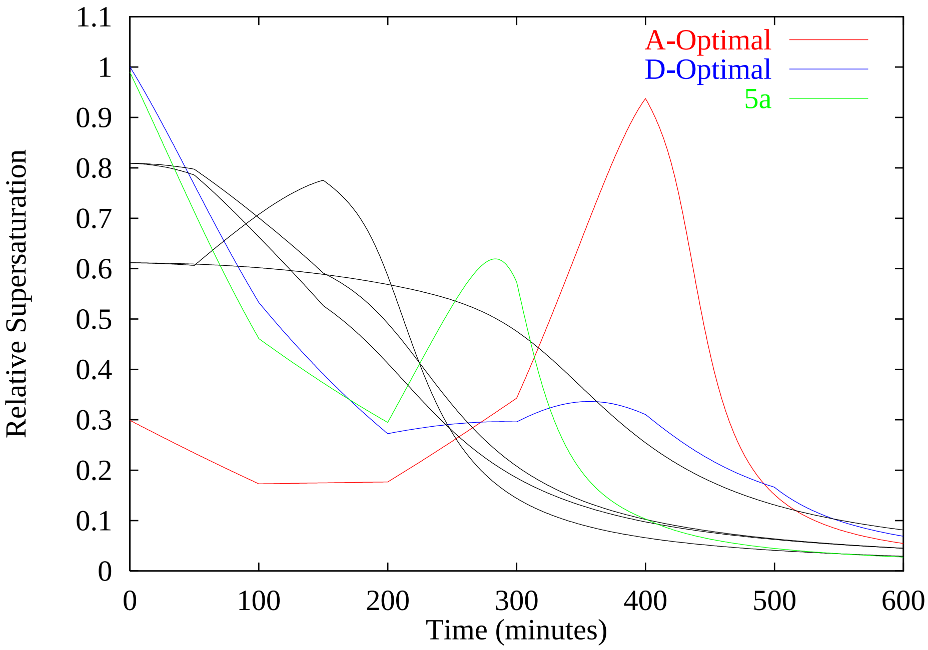

Figure 4.31: A-Optimal and D-Optimal relative supersaturation profiles as compared to initial profiles. Optimal profiles determined using the parameters in Table 4.4 with 0.25 gram seed mass.

unmapped regions of the space, the optimal designs are nearly orthogonal to the paths of previous relative supersaturation curves. This path minimizes the overlap of the information in the optimal experiment with that in previous experiments. The A-Optimal design is more distinct in this regard due largely to the similarities between the D-Optimal and run 5a profiles. The more distinct profile of the A-Optimal design is preferred because it extends the model over a wider, unmapped region of the operating space.

The D-Optimal and A-Optimal designs predicted mildly different sets of linear 95% confidence intervals. These intervals are shown with the prior intervals in Table 4.5. The D-Optimal design predicted greater reductions in the uncertainty of b, while the A-Optimal design was preferable for k_b and j. Reductions in the other parameter intervals were approximately equivalent for the two designs.

	$\ln(k_g)$	g	$\ln(k_b)$	b	j	$L_{\min}$	$l_w(\infty)$
Estimate	-1.50	1.89	29.57	3.72	2.21	28.13	17.37
Prior Int.	± 0.01	± 0.01	± 0.88	± 0.29	± 0.14	± 0.69	± 0.20
D-Optimal	± 0.01	± 0.01	± 0.87	± 0.23	± 0.13	± 0.57	± 0.19
A-Optimal	± 0.01	± 0.01	± 0.78	± 0.26	± 0.12	± 0.57	± 0.19

Table 4.5: Comparison of linear 95% confidence intervals from the revised model with the predicted intervals corresponding to the D-Optimal and A-Optimal designs.

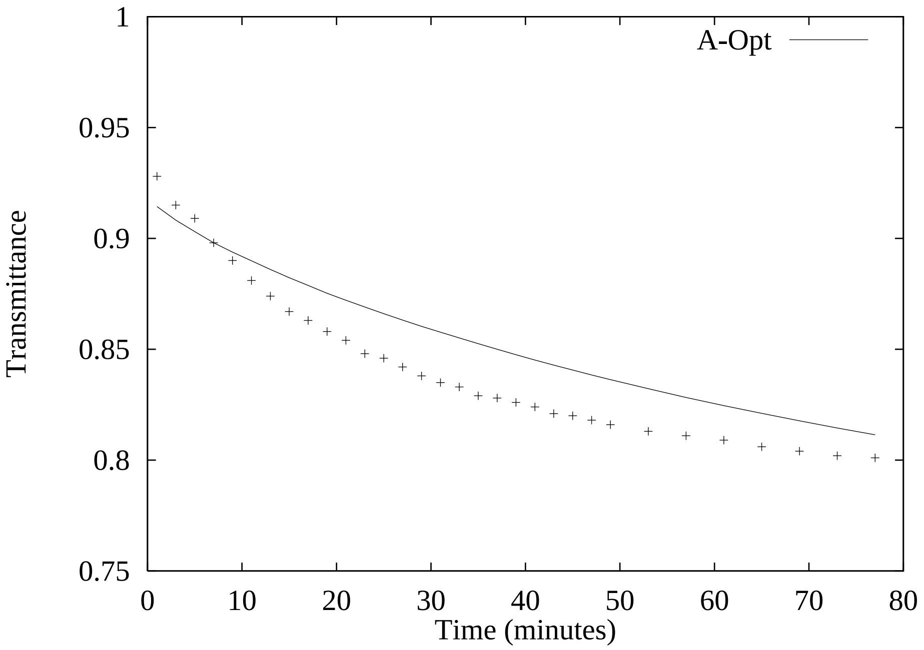

Figure 4.32: Transmittance profile and model fit for the A-Optimal experiment using the parameters in Table 4.6.

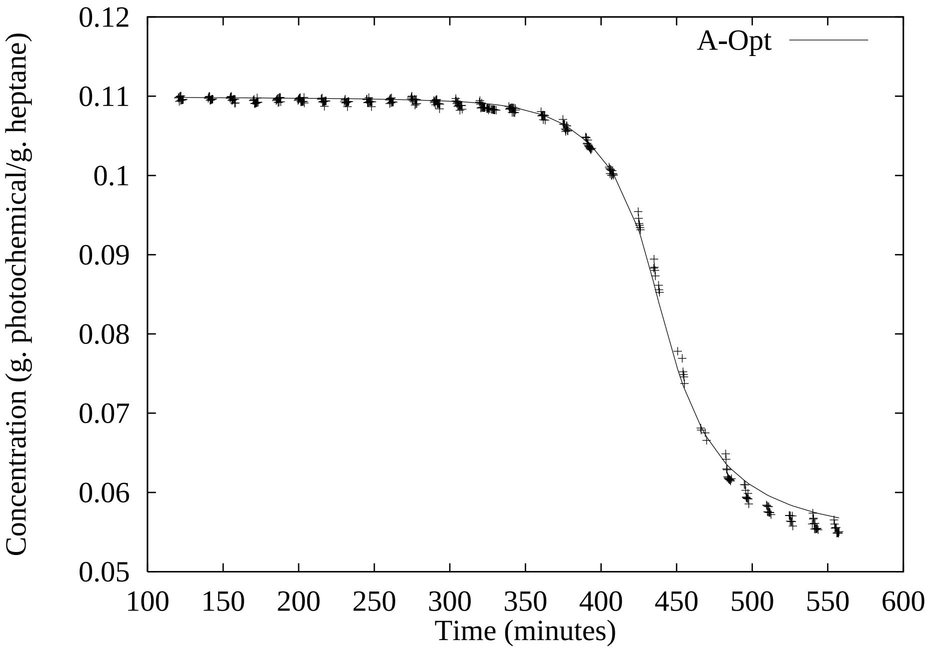

Figure 4.33: Concentration profile and model fit for the A-Optimal experiment using the parameters in Table 4.6.

In the absence of an advantageous reduction in parameter uncertainty, the A-Optimal design was chosen for the subsequent experiment because its relative supersaturation profile was the more unique of the two designs. Thus, it would test the model in a new region of the operating space. This experiment was conducted and the new transmittance and concentration data were appended to the previous data set. Given the additional data from the designed experiment, the model parameters were reestimated. These estimates are presented in Table 4.6. The actual confidence intervals of Table 4.6 closely approximate the predictions of Table 4.5. The improvements in the uncertainties of $ln(k_b)$, j, and $L_{\min}$ were essentially identical to the predicted values, while the improvements in $l_w(\infty)$ and j were better and worse than expected, respectively. The fit of the model to the data from the A-Optimal experiment is given in Figures 4.32 and 4.33. The fits of the new model to the prior data are indistinguishable from those depicted in Figures 4.23 through 4.28.

	$\ln(k_g)$	g	$\ln(k_b)$	b	j	$L_{\min}$	$l_w(\infty)$
Estimate	-1.53	1.88	29.53	3.76	2.17	28.08	18.24
Interval	± 0.01	± 0.01	± 0.80	± 0.28	± 0.12	± 0.58	± 0.16

Table 4.6: Posterior parameter estimates and 95% confidence intervals using concentration and transmittance data from the A-Optimal experiment.

No amount of data can prove a model to be true. Therefore, the decision to halt a model identification is a subjective one. In this study, the high quality of the predictions and the similarity between the new optimal parameter set and that estimated prior to the A-Optimal run were indications that the model was well-identified over the desired operating range. These facts, fortified by the more practical concern of the near exhaustion of the available supply of raw photochemical, motivated the decision to consider the body of collected data as adequate and to preserve the remaining supply of raw material for the optimal control phase of the project. However, this decision did not preclude additional analysis of the model based on the given data set. Thus, a series of model analysis and criticism issues are treated in the remaining sections of this chapter.

4.9 Microscopic Confirmation of Habit Dynamics

In Sections 4.6.2 and 4.6.1, the crystallizer model was revised with the addition of size-dependent nucleation and dynamic habit models. These revisions were justified subsequently by dramatic improvements in the fit of the model to transmittance data. However, no independent physical evidence was presented to support the revisions. The implementation of the A-Optimal design provided an opportunity to investigate these phenomena via optical microscopy.

During the A-Optimal experiment, one milliliter slurry samples were drawn from the crystallizer at 2, 18, 34, 299, and 349 minute time points. The samples were quenched in heptane and placed on microscope slides. The heptane was evaporated quickly with a compressed air stream, leaving the photochemical crystals isolated on the slides. A Nikon Optiphot-2 optical microscope equipped with a black and white electronic camera was used to examine and photograph the crystals. The photographs were downloaded onto a personal computer running Image Pro analysis software.

The photographs revealed that the crystal shapes were changing as they underwent growth. Figure 4.34 is a representative photograph of the seed crystals immediately after addition to the supersaturated solution. As indicated by the 50 micron scale, the crystals

	$\ln(k_g)$	g	$\ln(k_b)$	b	j	$L_{\min}$	$l_w(\infty)$
Estimate	-1.50	1.87	29.09	3.81	2.21	28.17	12.04
Interval	± 0.01	± 0.01	± 0.75	± 0.28	± 0.12	± 0.61	± 0.09

Table 4.7: Parameter estimates and 95% confidence intervals for concentration, transmittance, and optical microscope data using the revised model.

are small and have low length-to-width ratios. An analysis of this and other early-time photographs produced an initial length-to-width ratio estimate of 2.5.[4] The low length-to-width ratio of the seeds is contrasted by a representative photograph, Figure 4.35, taken after the seeds had grown several microns in width. The increase in the average length-to-width ratio is dramatic. This was predicted by the model as the optimal estimate of $l_w(\infty)$, 16.44 in Table 4.6, was significantly higher that the assumed initial value of 3.7.

In order to quantify the increasing length-to-width ratio as a function of ΔL, length-to-width ratios were measured from the optical micrographs using the Image Pro analysis software. These data were added to the concentration and transmittance data and the entire set was used to reestimate the parameters. The habit change model of Equation 4.31 was used to predict the length-to-width data. The optimal parameter estimates determined using the data from the micrographs are given in Table 4.7, with the length-to-width ratio fit given in Figure 4.36.

The addition of the microscope data lowered the optimal value of $l_w(\infty)$ to 12.04 from 18.24. However, the effect of the microscope data on the other estimates was minimal. The relative uncertainty of the linear confidence interval for $l_w(\infty)$ decreased to 7.5% of the parameter value from 8.7%.

[4]This contradicted the original estimate of 3.7 made from the SEM micrograph in Figure A.1. Because the newer, optical micrograph estimate was an *in situ* measurement, it was thereafter taken to be the appropriate estimate and was used in all subsequent work.

Figures 4.34 and 4.35 do not provide quantitative information about size-dependent nucleation. However, there is some indication in Figure 4.35 that the fast growing faces of the larger crystals are needle-like, perhaps due to dendritic growth. Since dendrites and needle-like structures are vulnerable to breakage [29], this may explain the importance of crystal-crystal collisions indicated by a third moment nucleation order approximately equal to two. The optimal value of $L_{\min}$ may indicate the average size required to remove dendritic structures in collisions.

4.10 Effect of the Data Set on the Estimates

It is clear from the Bayesian structure of the estimation problem that the modal estimates of the model parameters, $\hat{\boldsymbol{\theta}}$, are dependent on the content of the data matrix, $\boldsymbol{Y}$. Consequently, changes in $\boldsymbol{Y}$ translate into changes in $\hat{\boldsymbol{\theta}}$. One way to examine the robustness of a set of estimates is to look at how the point estimates and confidence intervals are affected by changes in the data set.

In this study, there was a concern that the experimental data collected prior to the identification of the densitometer-induced bias were not entirely compatible with data collected later in the study. Recall from the discussion in Section 4.5.3 that the discovery of the bias involved the removal of some data sets from $\boldsymbol{Y}$ but that much of the data were retained. The bias was removed in later runs by collecting concentration data intermittently. For convenience, the later runs (1c, 2c, 4c, and A-Optimal) are hereafter referred to as the unbiased runs while the early runs (2a, 2b, 3a, 3b, 4a, 4b, and 5a) are termed biased runs.

The intermittent collection of concentration data resulted in concentration data sets containing approximately 25% as many measurements as in the biased experiments. The difference is illustrated by Figure 4.28, where the replicate data sets from runs 4a and 4b are clearly more dense than that from 4c. If the optimal estimator is confronted with replicate data sets that are appreciably different, the reduction of the sum of squared prediction errors favors the estimation of parameters that reduce the errors in the larger set. In this case, such a bias in the estimates favored the biased runs that intuition suggested were the least likely to be correct.

In light of this concern, the parameter estimates presented earlier in this work were determined using all available concentration data from the unbiased experiments and every other (50% of available data) concentration data point from the biased sets. All available transmittance and length-to-width ratio data were used for all of the experiments. In order to ascertain the effect of leaving out data, the parameters were reestimated from two additional data sets. The first used all of the data from the four unbiased runs but used only every fourth (25% of available data) concentration data point from the early runs. This combination roughly equalized the the amount of concentration data taken from each experiment. A second estimation used all of the available concentration data from all of the runs in the study. This combination was appealing because it was not arbitrary. A comparison of the three sets of estimates is given in Table 4.8.

Conc. Data %	$\ln(k_g)$	g	$\ln(k_b)$	b	j	$L_{\min}$	$l_w(\infty)$
100	-1.48	1.94	31.89	4.15	2.64	27.20	12.01
50	-1.50	1.87	29.09	3.81	2.21	28.17	12.04
25	-1.52	1.82	27.11	3.56	1.90	29.06	11.91

Table 4.8: Parameter estimates as intermediate concentration data points from early, biased runs are excluded.

Table 4.8 shows clear differences in the optimal parameters estimated for each data set, indicating that the biased and unbiased data sets were not compatible. One way to alleviate such errors is to model the disturbance that caused the bias. As noted in Section 4.5.3, the disturbance involved the caking of crystals on the face of the densitometer filter. The effect was similar to a removal of crystals from the bulk. Caked crystals became immobile, biasing the transmittance measurement, and their surfaces were inaccessible to the solute molecules in the bulk solution, slowing the overall rate of mass transfer into the crystalline phase. This type of disturbance may be modeled by adding a stream term to the batch population balance of Equation 3.23. The new PBE is given as

$$\frac{\partial f(L,t)}{\partial t} = -G\frac{\partial f(L,t)}{\partial L} - k_f\left(\frac{q_d}{m_{sl}}\right) f(L,t) \tag{4.36}$$

where q_d is the mass flow rate in the densitometer sample stream, m_{sl} is the total mass of the slurry, and k_f is an adjustable parameter to be estimated from the data. Physically, k_f indicates how effectively the disturbance removed the crystals relative to a pure outlet stream with the same flow rate. In effect, Equation 4.36 models the caking as the transfer of well-distributed, kinetically active crystals into a subpopulation of immobile, inactive crystals. k_f is assumed to be zero for the unbiased experiments.

Using the disturbance model of Equation 4.36, the parameters were reestimated for

Conc. Data %	$\ln(k_g)$	g	$\ln(k_b)$	b	j	$L_{\min}$	$l_w(\infty)$	k_f
100	-1.39	1.76	26.94	2.91	2.20	27.08	11.18	0.39
50	-1.42	1.77	26.64	2.61	2.13	26.73	11.28	0.40
25	-1.48	1.77	26.19	2.30	2.01	25.91	11.91	0.41

Table 4.9: Parameter estimates for the disturbance model as intermediate concentration data points from early, biased runs are excluded.

the three data sets described above. The resulting estimates are summarized in Table 4.9. Though some variation remained in the optimal estimates, especially in the estimates of $\ln(k_g)$ and b, there was a marked improvement in the consistency of the estimate sets. The optimal estimates of k_f indicate that the disturbance removed crystals from the bulk slurry at a rate about 40% of that expected for a pure outlet stream with the same flow rate.

Adjusting for the densitometer disturbance led to a substantial change in the optimal estimates. However, the mechanistic interpretation of the estimates was unchanged with the estimates of g and j both close to 2 and the sign of $(b-g)$ remaining positive. This is a crucial point with regard to the optimal control strategies that are presented in Chapter 5. Despite the differences in the parameter estimates given in Table 4.8 and 4.9, the predicted optimal control strategies are nearly identical. This is demonstrated conclusively in Section 5.5.

Given the consistency of the disturbance model estimates and the desire to avoid the arbitrary removal of data, the optimal model and parameter set was taken to be the 100% concentration model in the first row of Table 4.9. The parameters and confidence intervals for this model are given separately in Table 4.10 and the residual fits to the data are given in Figures 4.37 through 4.42.

	$\ln(k_g)$	g	$\ln(k_b)$	b	j	$L_{\min}$	$l_w(\infty)$	k_f
Estimate	-1.39	1.76	26.94	2.91	2.20	27.08	11.18	0.39
Interval	±0.01	±0.02	±0.32	±0.16	±0.05	±0.30	±0.17	±0.02

Table 4.10: Parameter estimates and 95% confidence intervals for the optimal disturbance model using all available concentration, transmittance, and optical microscope data.

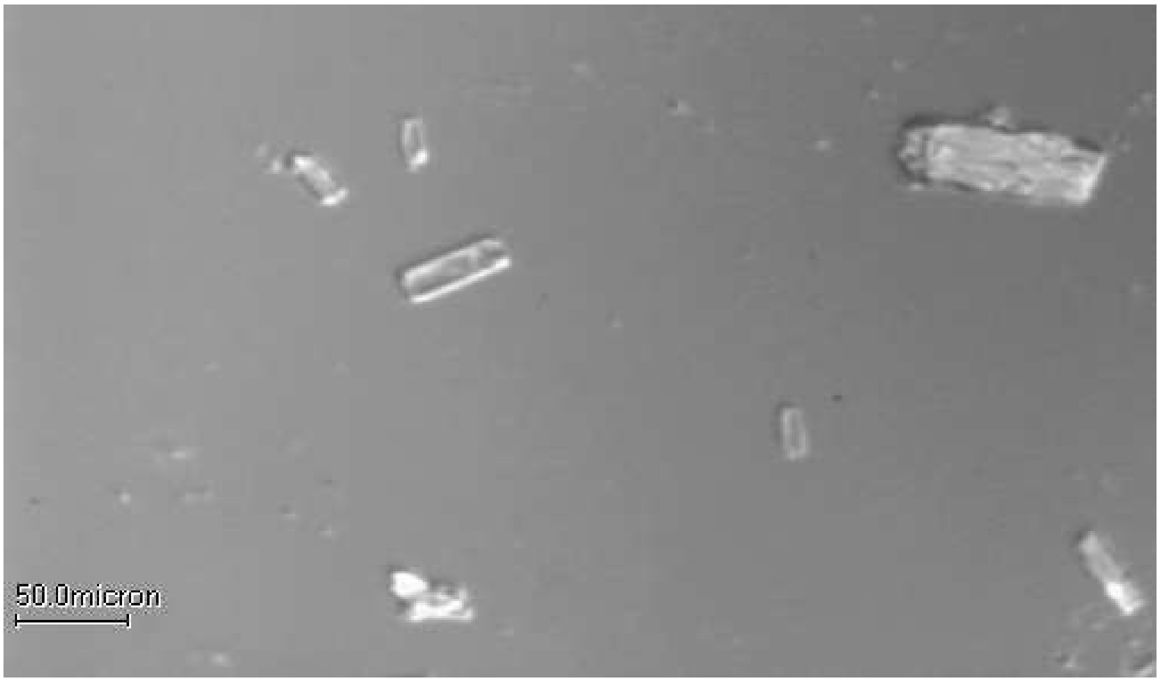

Figure 4.34: Optical micrograph of low length-to-width ratio seed crystals.

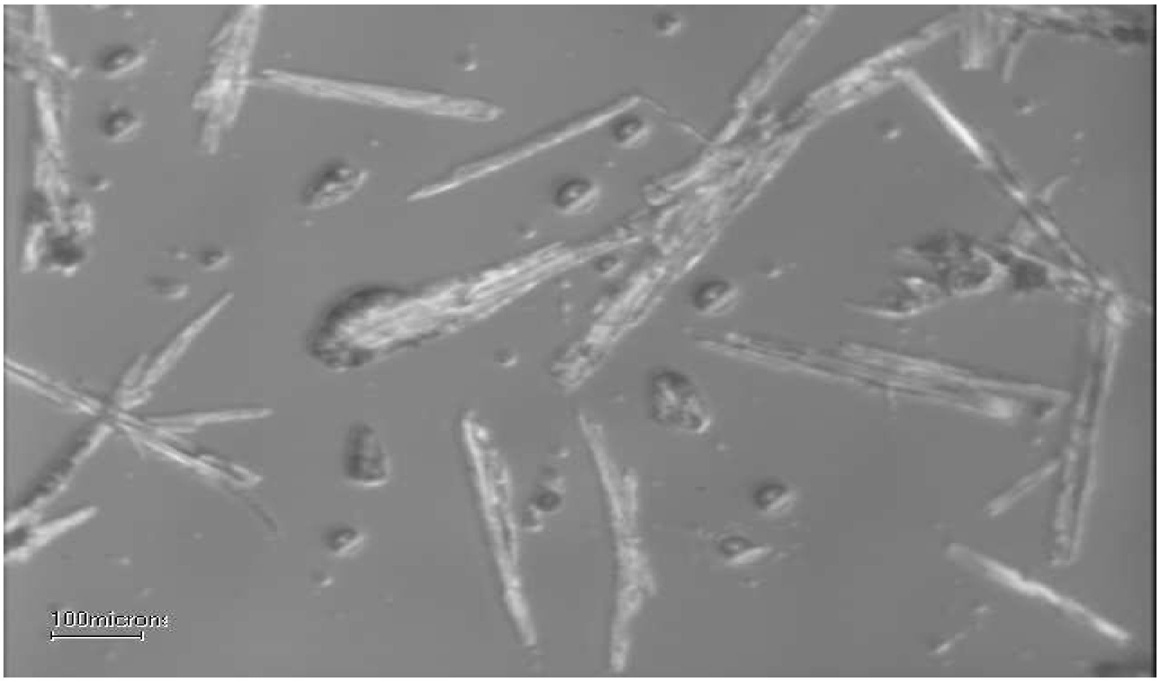

Figure 4.35: Optical micrograph of high length-to-width ratio crystals later in same crystallization.

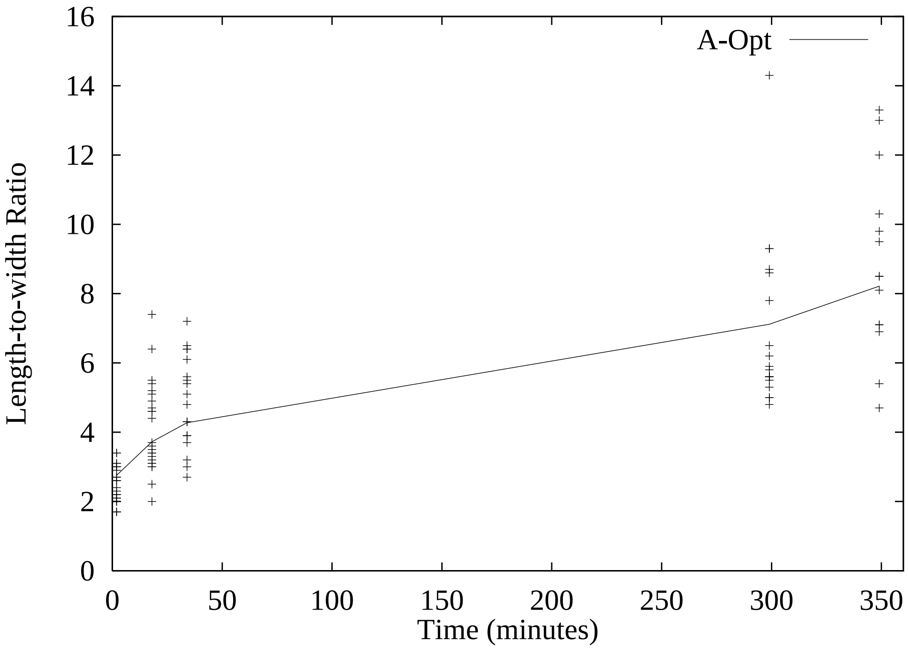

Figure 4.36: Length-to-width ratio versus time profile for the A-Optimal design.

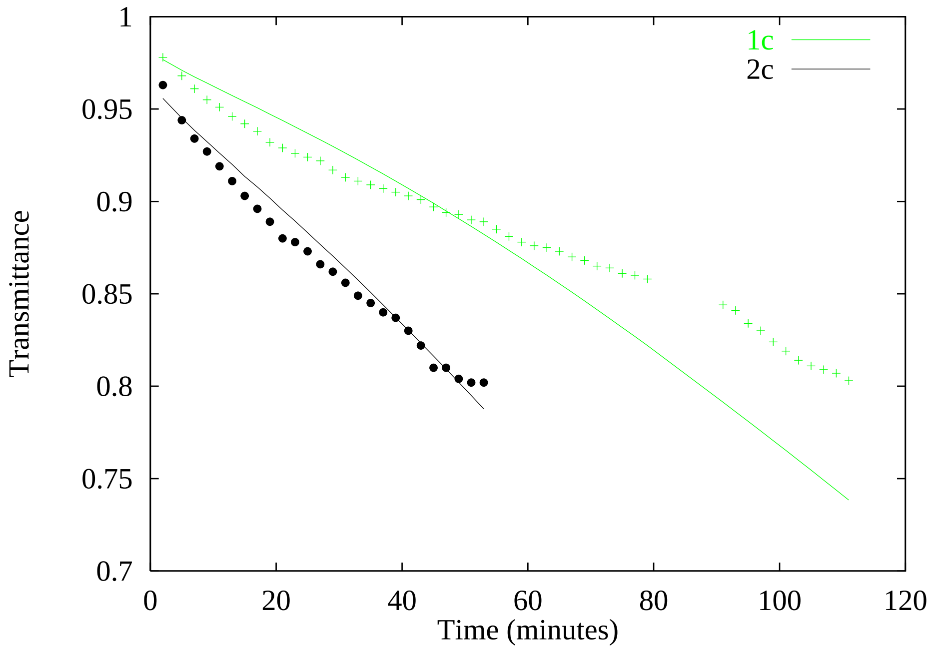

Figure 4.37: Transmittance profiles and model fits for runs 1c and 2c using the parameters in Table 4.10.

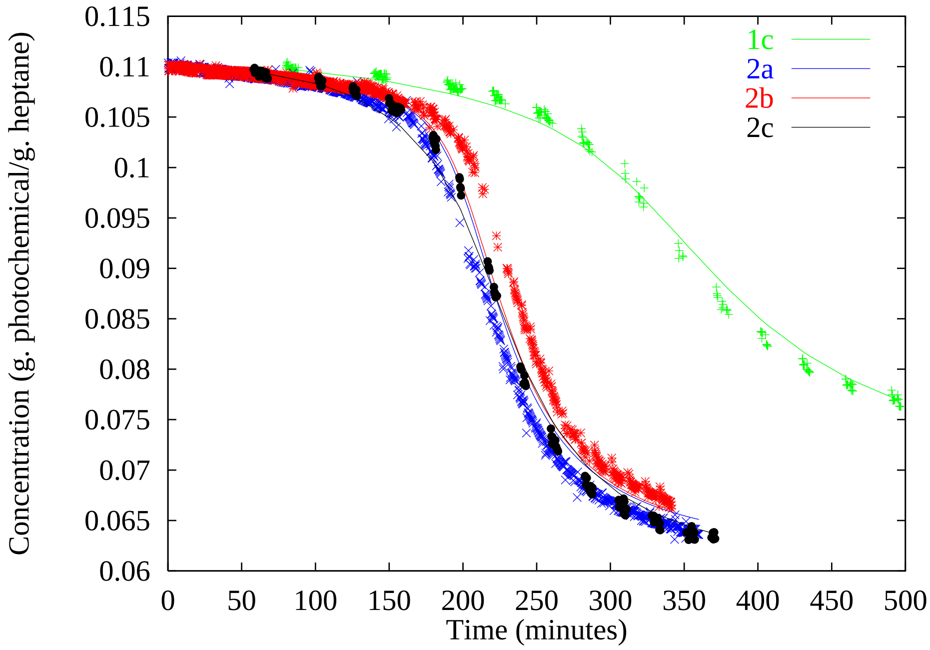

Figure 4.38: Concentration profiles and model fits for runs 1c, 2a, 2b, and 2c using the parameters in Table 4.10.

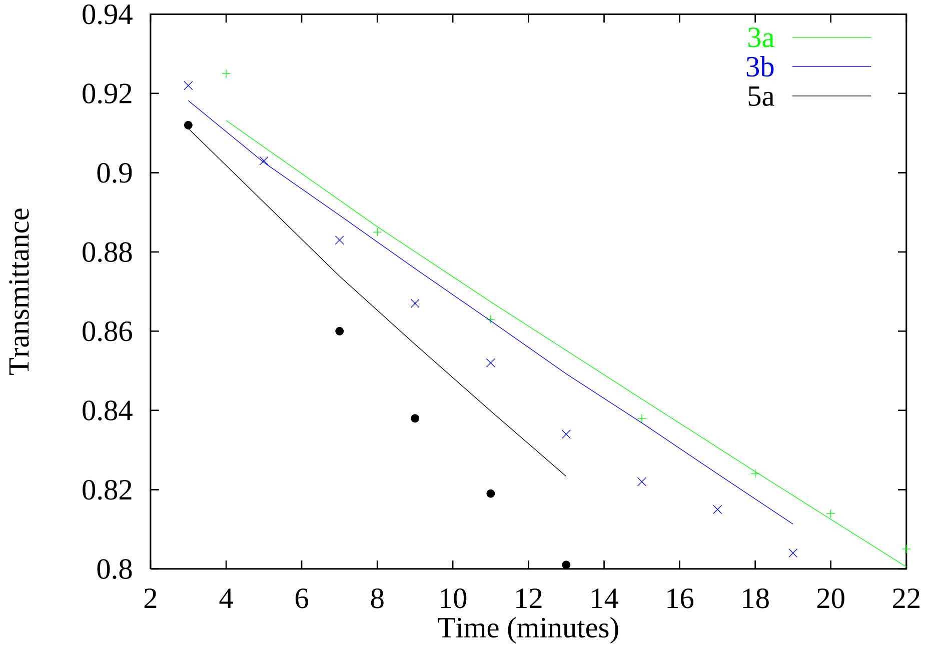

Figure 4.39: Transmittance profiles and model fits for runs 3a, 3b, and 5a using the parameters in Table 4.10.

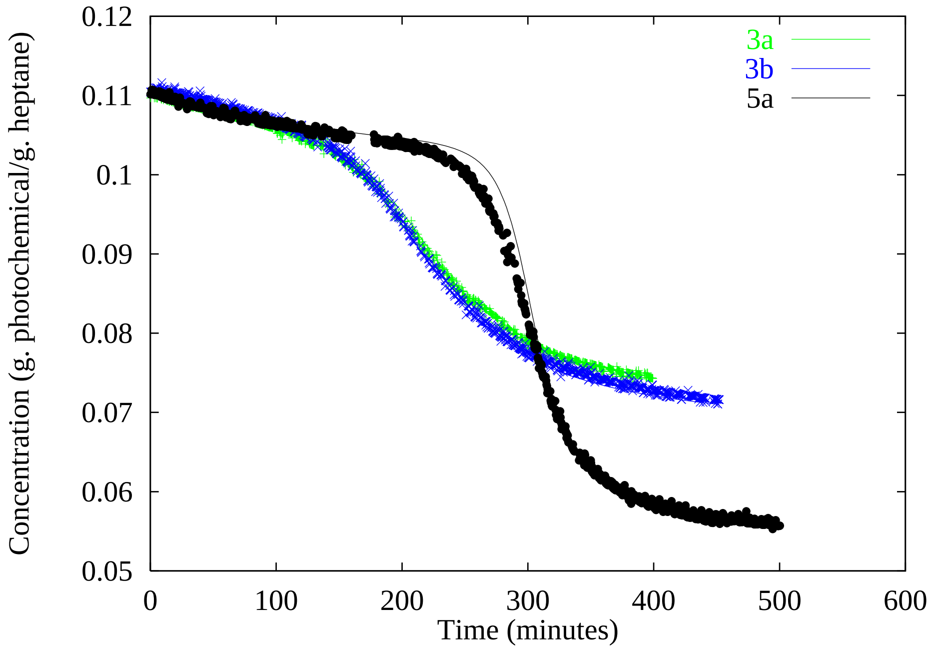

Figure 4.40: Concentration profiles and model fits for runs 3a, 3b, and 5a using the parameters in Table 4.10.

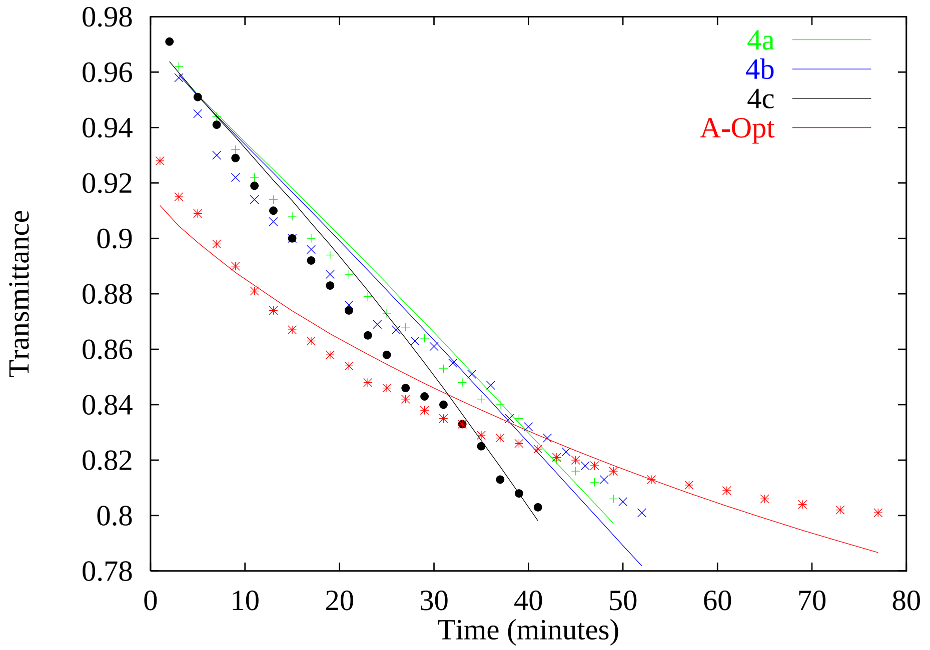

Figure 4.41: Transmittance profiles and model fits for the A-Optimal design and runs 4a, 4b, and 4c using the parameters in Table 4.10.

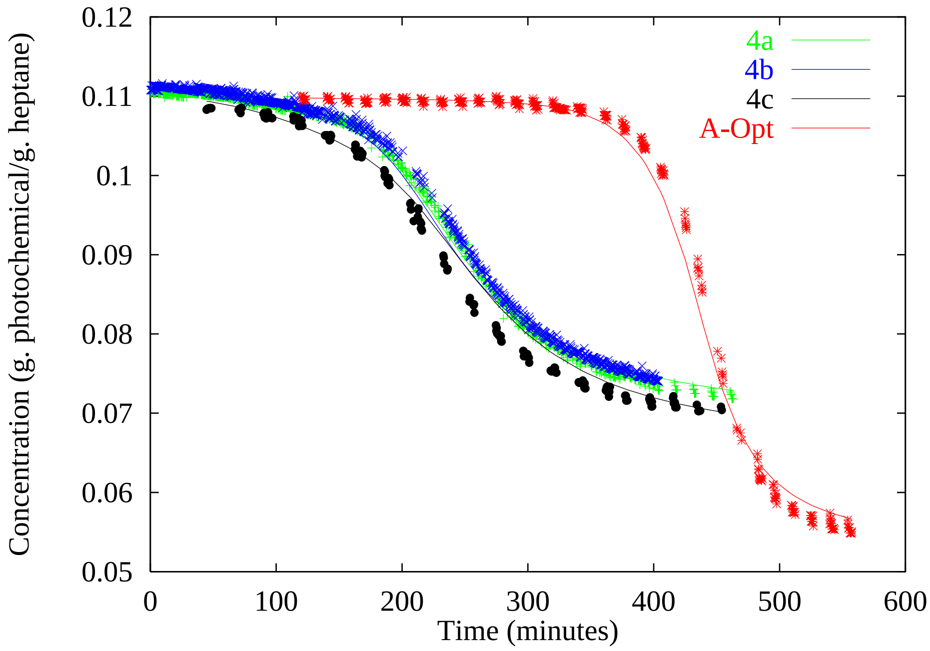

Figure 4.42: Concentration profiles and model fits for the A-Optimal design and runs 4a, 4b, and 4c using the parameters in Table 4.10.

4.11 Analysis of Parameter Convergence

The estimation of parameters via nonlinear optimization treats the problem of parameter observability through the convergence of the nonlinear optimizer. For an unconstrained optimization, convergence implies that at a local optimum the gradient of the objective function with respect to the parameters is zero and that the Hessian is positive definite. If the objective function surface is flat with respect to a subset of the parameters, these criteria may not be met and the lack of convergence indicates that the information in the data set is insufficient to allow estimation of those particular parameters. However, in practice optimizers often have poor gradient information near the optimum. Poor gradient information prevents the optimizer from determining search directions that minimize the objective function. It also prevents the verification of the optimality of a current iteration of the parameters. When the optimizer does not converge explicitly the objective function surface must be investigated to determine whether or not an optimum has been achieved.

For the parameters given in Table 4.10 the optimizer did not converge. Therefore, the nonlinear objective function surfaces of the parameters were examined. It was determined from these plots that the noise in the data was preventing convergence by roughening the surface near the optimum. Figure 4.43 illustrates the roughening by plotting the objective function value as the parameter $l_w(\infty)$ is varied about its determined value. The other parameters were held at the values given in Table 4.10. Note the jagged path of the objective function near the optimum. The uneven surface is more apparent with respect to $l_w(\infty)$ than with any other parameter. This reflects the fact that the noisy microscope data shown in Figure 4.36 largely determine the optimal value of $l_w(\infty)$. Near the optimum, small changes in $l_w(\infty)$ alter the large prediction errors of the microscopy data, minimizing some errors while others grow larger. The effect is a corrugated objective function surface.

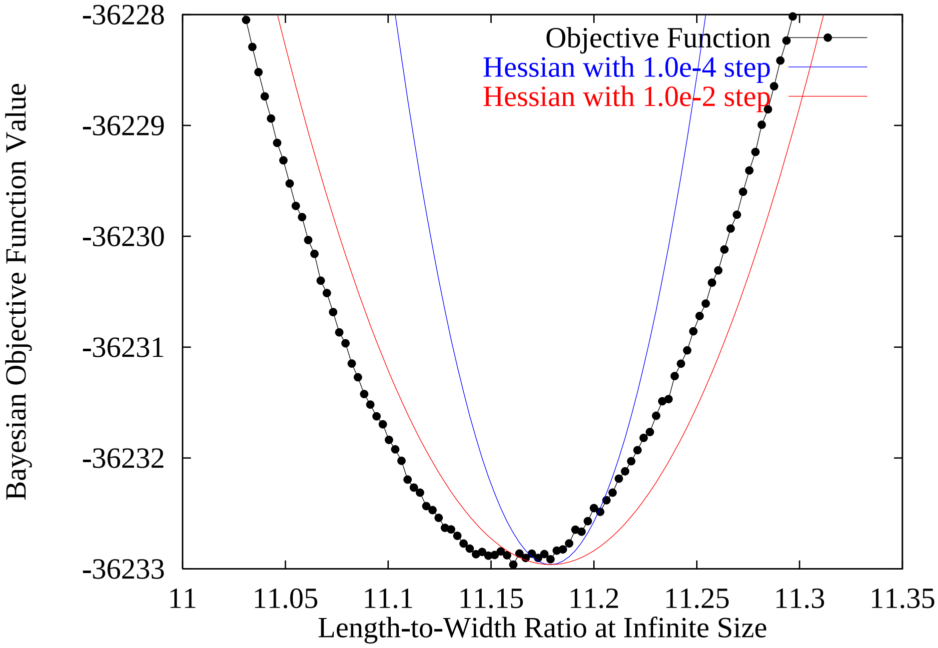

Figure 4.43: The change in the objective function as $l_w(\infty)$ is varied about its value of 11.18. The smooth curves represent the surface predictions for linear finite-difference Hessians determined with different step sizes.

Showing that the objective function surface is non-smooth helps to explain the lack of convergence but it does not provide proof of a local optimum. However, if the solution is locally optimal, each of the individual parameters must be at a minimum in its own coordinate plane. This may be shown by examining objective function plots similar to Figure 4.43 for each parameter. For the sake of brevity, these plots are not provided in this thesis, however, in each case the values appeared to be at minima. Given this limited evidence, the parameter values of Table 4.10 were assumed to represent a local minimum.

As an aside, the construction of objective surface plots for each parameter allows a simple check on the adequacy of the linear Hessian. For the ith parameter, the i, i element of the Hessian may be used to predict the shape of the objective function near the optimum. This is illustrated for $l_w(\infty)$ in Figure 4.43. The two Hessian predictions correspond to linear Hessians generated with finite-difference step sizes of 1.0×10^{-4} and 1.0×10^{-2}. Despite the shifted value of the minimum, the larger step size clearly provides a better prediction of the surface. Similar step-size effects were observed for the other parameters as well. The effect of step size on the summed squared errors of the Hessian predictions of the objective function as the parameter g is varied are summarized in Figure 4.44. Again, the larger step sizes appeared to provide improved fits to the surface.

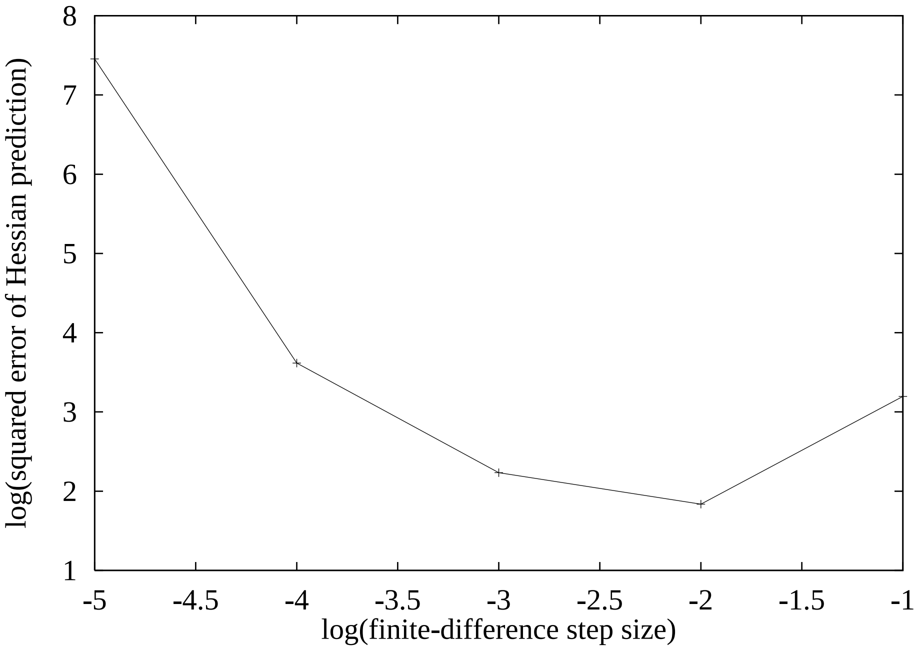

Figure 4.44: The change in the summed squared errors for linear Hessian predictions of the objective function as g is varied about a value of 1.76.

Since the original Hessian had been estimated using 1.0×10^{-4} step size, a reestimation of the intervals was required. The linear confidence intervals for the best parameter set were reestimated using a step size of 1.0×10^{-4}. The parameters and new intervals are given in Table 4.11. The new intervals are considerably larger, reflecting the flatter objective function surface. Additional examinations of the linear Hessian assumption are given in the following section.

	$\ln(k_g)$	g	$\ln(k_b)$	b	j	$L_{\min}$	$l_w(\infty)$	k_f
Estimate	-1.39	1.76	26.94	2.91	2.20	27.08	11.18	0.39
Interval	± 0.04	± 0.02	± 0.57	± 0.18	± 0.08	± 0.67	± 0.81	± 0.03

Table 4.11: Parameter estimates and 95% confidence intervals for the optimal disturbance model using all available concentration, transmittance, and optical microscope data.

4.12 Analysis of Serially Correlated Errors and Linearity

The estimates and confidence intervals of the preceding sections were determined subject to assumptions concerning the probabilistic structure of the errors and the linearity of the posterior density function about its mode. In particular, the prediction errors of the same type were assumed to be serially independent. This assumption is examined via the residuals corresponding to the optimal parameter estimates of Table 4.11. The appropriateness of the linear Hessian is checked in two ways. First, by evaluating the nonlinear objective function along the elliptical boundary of the linear HPD region and second by examining the relative posterior density within the linear confidence interval bounds.

4.12.1 Examination of Serial Independence

Serial correlation, also termed autocorrelation, is the correlation of errors in time that may occur due to unmodeled dynamics or disturbances in the data. It is especially common when, as in this study, relatively precise sensors and short sampling times are used to record data in dynamic experiments [15]. Representative examples of serial correlation in the concentration residuals of this study are illustrated for runs 2a and 2b in Figure 4.45.

The residuals in Figure 4.45 are characterized by three time periods. The first period accounts for the first 150 minutes of the experiments and indicates the accuracy of the initial condition. During this period the prediction errors are relatively small for both sets. The second period lasts from 150 to approximately 300 minutes and is characterized by large positive and negative residuals. Recalling the run 2 data from Figure 4.38, these residuals correspond to a period of strong concentration gradients in which small errors in the initiation time of the high gradient period are translated into large prediction errors. At the end of the run, a third period of small residuals illustrates the replicability of the solubility curve as the experiments approach equilibrium. In each period, the residuals are serially correlated.

In some cases, serially correlated residuals may indicate that there are significant structural errors in the model. However, the serial correlations of Figure 4.45 are more indicative of the lack of replicability of the data. According to the model, the conditions of runs 2a and 2b are essentially identical, therefore the model predicts nearly identical concentration profiles for the two runs. However, the differences in the concentration data sets collected for runs 2a and 2b are large relative to the precision of the concentration sensor. Under these conditions, the model cannot accurately predict both data sets and serial correlations always occur. The key issue is the high precision of the measurement sensor relative to the replicability of the data. The fact that the errors are positive for one replicate and negative for the other, suggests that the optimal parameter set minimized the errors by selecting a middle path.[5] The positive and negative errors are favorable outcome indicating that the serial correlations are more a function of data replicability than of structural bias in the model.

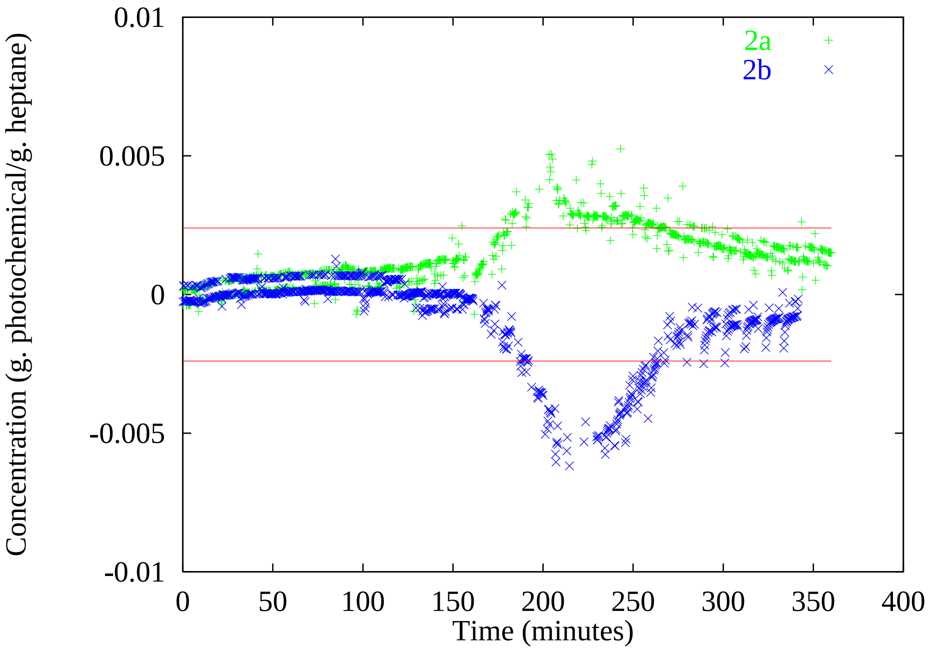

Figure 4.45: Concentration residuals for the run 4 replicates using the parameters of Table 4.11. Solid lines enclose a $\pm 2\hat{\sigma}_c$ region based on Equation 4.20.

The transmittance residuals for the run 4 replicates are shown in Figure 4.46. In general, the upward concavity indicated in Figure 4.46 is preserved for all of the data sets. Thus, the transmittance residuals indicate more significant structural biases than did the concentration residuals. The existence of structural errors in the transmittance model is not surprising given the sensitivity of the transmittance measurement to crystal habit and seed size distribution. Given the gross assumptions (i.e. parallelepiped shape, average length-to-width ratio, linear seed size distribution) made to define habit and seed distribution

[5]Similar trade-offs occur for all of the concentration residual sets in the study. The run 2a and 2b residuals were chosen because they illustrated most clearly the replicability of the data relative to the measurement precision.

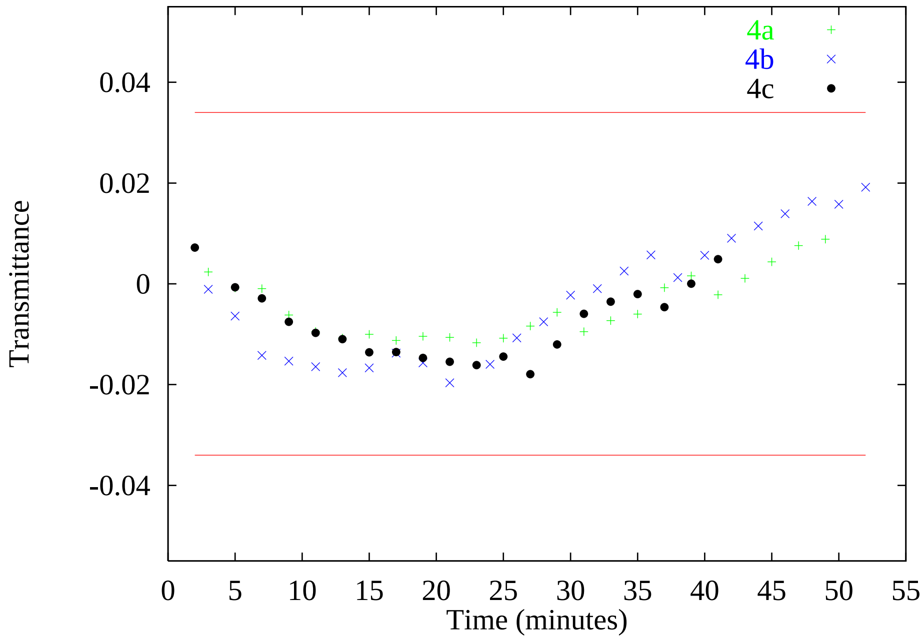

Figure 4.46: Transmittance residuals for the run 4 replicates using the parameters of Table 4.11. Solid lines enclose a $\pm 2\hat{\sigma}_{tr}$ region based on Equation 4.20.

in the model, the existence of structural error is not surprising. However, the prospects of accounting for structural errors are slim without an improved knowledge of the crystal habit or seed size distribution.

4.12.2 Linearity of the Posterior Density

It was assumed that the nonlinear posterior density function could be approximated by linear functions near its mode. The linear assumption allowed the calculation of a linear Hessian and the subsequent determination of HPD regions and 95% linear confidence intervals for the parameters. If the linear assumption is applicable, these boundaries are hyperellipsoids of constant probability density. One way to determine the adequacy of this assumption is to examine the uniformity of the nonlinear objective function values along the boundaries of the HPD region. Bard [4] suggests evaluating the posterior density at the endpoints of the axes of the HPD hyperellipsoid. As illustrated in Figure 4.29, these endpoints are defined by eigenvectors and eigenvalues of $\boldsymbol{V}_\theta$.

With 8 estimated parameters in the optimal model, the objective function may be evaluated at 16 points along the HPD region boundary. For each axis, evaluations are made in the positive and negative eigenvector directions. Table 4.12 gives the difference between the optimal value of Φ_{be} and those measured at the endpoints of the axes.

	1	2	3	4	5	6	7	8
+	15.5	14.3	14.4	13.5	11.8	18.0	78.8	4.6
−	15.5	15.6	14.2	17.0	14.8	17.9	92.5	20.3

Table 4.12: Difference between $\Phi_{\text{be}}(\hat{\boldsymbol{\theta}})$ and $\Phi_{\text{be}}(\boldsymbol{\theta})$ evaluated at endpoints of the axes that define the 95% HPD region. The linear Hessian predicts a difference of 15.5.

The linear Hessian gives an approximation for the objective function value along the HPD region boundary

$$\hat{\Phi}_{\text{be}} + \frac{1}{2}(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})\boldsymbol{H}_{\hat{\theta}}(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}). \tag{4.37}$$

For the optimal disturbance model, the difference between the linear approximation and the optimal solution is 15.5. The linear approximation appears to provide a good approximation for the first 6 of the 8 axes. However, there is clearly variation among the values of Table 4.12. Of particular concern is the assymetry along the eighth axis, which may be caused by a poorly determined optimum along this axis or may be the result of the nonlinearity of the model.

In systems of two or three parameters, a convenient way to more closely examine the linear assumption is to plot the contours of the objective function surface. With 8 parameters, this is difficult to present clearly. However, it is reasonable to plot the objective function as two parameters are varied within their confidence intervals and the rest of the parameters are held at their modal values. Figure 4.47 plots the difference between $\Phi_{\text{be}}(\hat{\boldsymbol{\theta}})$ and $\Phi_{\text{be}}(\boldsymbol{\theta})$ as the supersaturation orders g and b are varied within their 95% confidence limits. The contour line indicates the level of the linear correction, 15.5.

The information garnered from Figure 4.47 is largely qualitative. It is comforting that the nonlinear surface is convex with respect to g and b and that their optimal values appear to locate the center of the elliptical contour. Note that the elliptical contour corresponding to the linear approximation is entirely contained within the rectangular plane of the 95% confidence region. This implies that the linear 95% confidence intervals are reasonable for this parameter pairing. However, it must be recognized that containing the linear contour for the joint distribution of two parameters is only a necessary condition for the linear assumption to apply. It is sufficient only if the parameter coordinates are also the axes of the hyperellipse (i.e. the parameters are uncorrelated) and if the contour is likewise contained for all other combinations of the parameters.

The information in Figure 4.47 may be presented in terms of posterior density functions by plotting the ratio $p(\boldsymbol{\theta}|\boldsymbol{Y})/p(\hat{\boldsymbol{\theta}}|\boldsymbol{Y})$ versus the varied parameters. This ratio is defined by the equation

$$p(\boldsymbol{\theta}|\boldsymbol{Y})/p(\hat{\boldsymbol{\theta}}|\boldsymbol{Y}) = \exp\Big((-\Phi_{\text{be}}(\boldsymbol{\theta})) - (-\Phi_{\text{be}}(\hat{\boldsymbol{\theta}}))\Big)$$

where the sign change is necessary to account for the fact that the minimization of the objective functions corresponds to the maximization of the posterior density. Figures 4.48 and 4.49 illustrate such plots for the b versus g and $\ln(k_b)$ versus j parameter combinations. The $\ln(k_b)/j$ combination was investigated because the linear Hessian predicted a strong

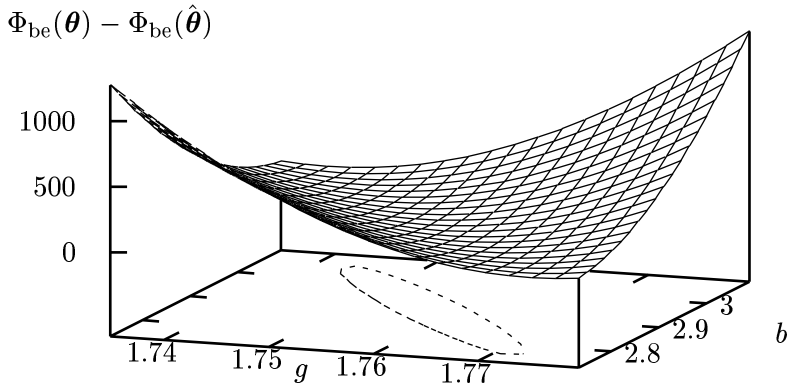

Figure 4.47: Objective function surface as g and b are varied within their respective 95% confidence limits. Contour line indicates the linear correction value, 15.5, predicted by Equation 4.37. Other parameters are held at the nominal values of Table 4.11.

correlation for these parameters. This correlation is shown in Figure 4.49. In each figure the linear 95% confidence intervals enclose fully the regions of appreciable probability. Again, this does not guarantee the sufficiency of the linear intervals. However, it does imply that the conservative linear intervals defined by the hyperbox are adequate if the remaining parameters reside near their modal values.

As noted by Stewart et al. [141], the explicit integration of the posterior density function is difficult. Therefore, the assumption of linearity about the mode of the posterior density, or the natural logarithm of that density, is adopted as a convenient way to convey information about relative degrees of uncertainty among the elements of $\hat{\boldsymbol{\theta}}$. When, as in this case, the elliptical HPD region does not agree perfectly with the shape of the posterior density, the insufficiency of the linearity assumption must be noted so that it is clear that the cited intervals may not correspond to the precise confidence implied by the choice of

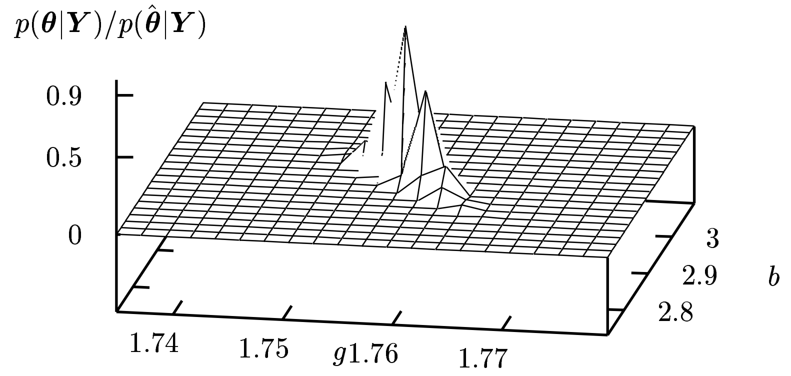

Figure 4.48: Relative posterior density surface as g and b are varied within their respective 95% confidence limits. Other parameters are held at the nominal values of Table 4.11.

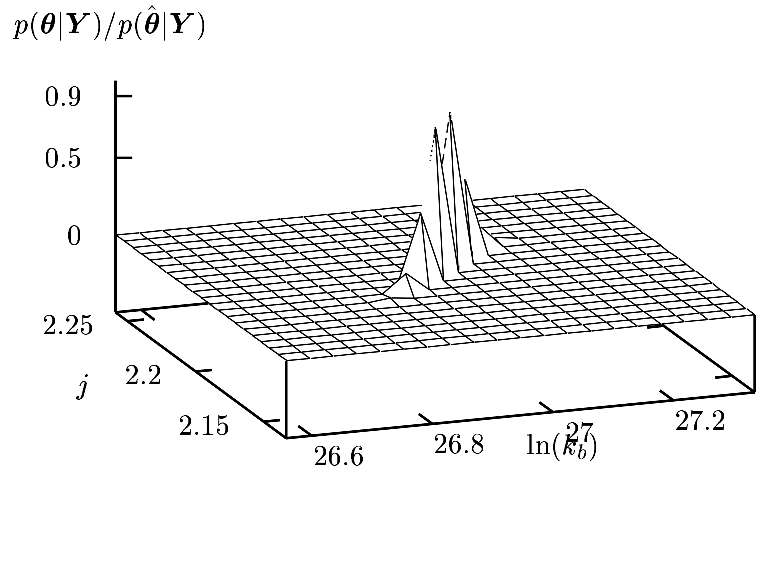

Figure 4.49: Objective function surface as $\ln(k_b)$ and j are varied within their respective 95% confidence limits. Other parameters are held at the nominal values of Table 4.11.

the $\chi_p^2(\alpha)$ bound. However, especially in light of favorable checks on the nonlinear objective or posterior density surfaces, the inaccuracy of the linear assumption does not necessarily imply that the linear intervals are without merit. The intervals still provide a measure of the relative uncertainty of the estimates and may represent approximately the regions of significant posterior density for the parameters.

4.13 Summary

This chapter has presented the results of a detailed model identification study for the kinetic processes of crystal nucleation and growth in the photochemical-heptane system. The kinetic parameters were estimated using liquid phase concentration and slurry transmittance data collected on-line in the batch crystallization apparatus described in Chapter 2. The study followed an iterative progression from an insufficient initial model through revisions in data collection and model structure to a final set of optimal parameter estimates for the revised model. The revised structure of the general model included the addition of a dynamic crystal habit model and a size-dependent crystal nucleation model. The dynamics of the crystal habit were confirmed experimentally using optical microscopy. Optimal experimental design methods were presented and applied to design crystallization experiments to reduce the uncertainty in the optimal parameter estimates.

The parameter estimates in the revised model were analyzed with respect to different data sets. These data sets tested the compatibility of data collected from the 11 experiments used in the estimation. A disturbance model was proposed and identified to reduce the effect of a bias that affected some of the data sets. The optimal parameter estimates and 95% confidence intervals determined using this disturbance model are given in Table 4.11. Additional analyses concerning the importance of the transmittance data in the estimation of the parameters and the failure to estimate dissolution parameters are given in Appendix F.

Chapter 5 uses the optimal disturbance model identified in this chapter to determine optimal open-loop control strategies for the batch cooling crystallizer. These strategies are designed to improve the filtration characteristics of the final-time crystallization slurries.

Chapter 5

Control of Batch Crystallizers

Once the model and kinetic parameters describing a particular process have been identified, a logical next step is to use the model to optimize the performance of the process. In batch processes this involves the determination of open-loop manipulated variable profiles that optimize a chosen performance measure. Optimal control theory provides tools that allow the determination of such profiles subject to the identified model and realistic constraints on the inputs, outputs, and paths of the states.

As stated in the introduction, a goal of this study was to improve the filtration properties of the photochemical/heptane system. This chapter describes the accomplishment of this goal through the application optimal temperature profiles. The profiles were optimized by minimizing the nucleated crystal mass relative to the mass of grown seeds $(m_N(t_f)/m_S(t_f))$. Improvements in filtration performance were verified by comparing the specific cake resistance values of filter cakes produced by optimized temperature profiles to those corresponding to the identification experiments of Chapter 4.

Chapter Overview. The chapter begins in Section 5.1 with a review of previous work performed in batch crystallization control. The focus is on the determination of optimal open-loop temperature profiles, but feedback control efforts are also noted. Section 5.2 presents the nonlinear optimization method used to determine constrained optimal temperature profiles for the photochemical-heptane system. The initial conditions, constraints, and temperature profile parameterization required to initialize the control problem are presented Section 5.2.2. Section 5.3 describes optimal temperature profiles maximizing final-time seed size and minimizing $m_N(t_f)/m_S(t_f)$. Section 5.4 examines the effects of seed mass, run duration, and the supersaturation orders of nucleation and growth on the m_N/m_S-optimal profile. The sensitivity of the m_N/m_S-optimal profile to uncertainty in the model parameters is analyzed in Section 5.5. Section 5.6 describes the implementation of two optimal temperature profiles and comments on a shift in the kinetics of the photochemical-heptane system. Section 5.7 compares the filtration results of the optimal experiments to previous filtrations performed for the model identification experiments described in Chapter 4.

5.1 Optimal Control in Batch Crystallization

In batch crystallization the optimized manipulated variable is typically the one that controls the nonequilibrium driving force in the system. In precipitation or reactive crystallization processes this variable may the addition rate of a nonsolvent or a reactant. In batch cooling crystallization the temperature of the slurry is the variable of interest. The concept of controlling cooling is introduced by Griffiths [59]. He proposes maintaining supersaturation levels below the meta-stable limit to prevent the spontaneous production small crystals due to primary nucleation mechanisms [154]. Though not strictly applicable to seeded crystallization systems that do not experience primary nucleation, Griffiths' idea of minimizing nucleation remains the basic tenant of batch crystallization control.

Mullin and Nyvlt [99] investigate controlled cooling in aqueous Potassium and Ammonium Sulfate systems by comparing the size and uniformity of the crystals produced by a controlled profile with those from a natural cooling (i.e. maximum cooling) experiment. Using a simple mass balance model, the control profile is designed to maintain a constant level of supersaturation. The level is presumed to be low enough that nucleation does not occur. The authors verify that the constant supersaturation profile produces larger crystals and a more uniform distribution than does the natural cooling experiment. The optimal cooling curves follow a pattern of slow initial cooling followed by higher cooling rates at the end of the crystallization. Ulrich [147] shows similar benefits for an organic product in an evaporative crystallizer. However, he arrives at his optimum cooling empirically by examining sieved product samples for a variety of cooling curves. The general shape of these curves was based on the optimal shape advocated by Mullin and Nyvlt [99].

In a subsequent simulation study, Ajinkya and Ray [2] develop an optimal operating policy using the Ammonium Sulfate model of Mullin and Nyvlt [99]. They pose the problem as the maximization of average crystal size with respect to the temperature profile followed during the crystallization. The results show that even when nucleation kinetics are neglected, the intuitive strategy of maximizing cooling rate in order to maximize growth rate may not be an optimal policy. This is due to the positive correlation between temperature and the growth rate constant. Essentially, the low temperatures required to elevate driving force levels also lower the growth rate constant. Morari [96] develops an analytical solution for this formulation of the batch optimal control problem. These two studies provide valuable insights into the optimal cooling policy problem but are limited by the neglect of nucleation kinetics and lack of experimental confirmation.

Using a method similar to that of Ajinkya and Ray, Jones [80] optimizes batch cooling profiles using moment formulations of the general, size-independent population balance model. Jones optimizes terminal seed size and confirms experimentally that the optimal cooling trajectory increases average crystal size as compared with linear and natural cooling strategies. Chang and Epstein [25] use simulations of a moment-based PBE model to examine a feedback strategy for maintaining previously optimized profiles. Their strategy uses a one-step iteration to improve the final-time performance index for a perturbed state profile. The assumption is made that the states of the perturbed profile may be described by a model linearized about the states of the desired optimal profile. The benefits of this scheme are not verified experimentally.

Miller and Rawlings [94] develop a more flexible optimal control scheme that allows the use of the full nonlinear model and the application of constraints. Their method is, with the exception of the objective function, identical to that used to determine the optimal experimental designs described in Section 4.8.3 of this thesis. This technique involves the optimization of a piecewise linear cooling curve with respect to some function of the final-time moments of the CSD. Unlike previous techniques, the formulation allows straight forward incorporation of cooling rate, state, and final time constraints. This flexibility allows the method to select policies that meet yield constraints without violating supersaturation bounds or exceeding available cooling capacity. Using kinetic parameters determined for the KNO_3–H_2O system, Miller and Rawlings [94] implement an optimal strategy minimizing the ratio of nucleated seed mass to grown seed mass. Their experimental results indicate average weight-mean crystal size increases of 28% percent versus natural cooling and 2%

versus linear cooling. The process yield was not diminished by the optimal policy.

More conventional feedback control work is performed by Rohani et al. [130]. These authors measure fines density and calculate the appropriate fines dissolution flow rate required to return the system to a prescribed fines density profile. The nominal policy is not optimized but the fines density is controlled, promoting the growth of larger crystals by reducing the crystal surface area of the competing class of nuclei. Experimental implementation on a potash alum-water system results in final mean crystal size improvements of up to 80% for linear and natural cooling runs. However, the use of fines dissolving diminished final time yield by up to 35%. Farrell and Tsai [40] successfully employ a similar program tracking a weight mean size profile measured by a light scattering device.

Févotte and Klein [42] take an innovative approach to feedback control of vacuum cooled batch crystallizers. The problem is to follow a desired cooling curve without requiring heating loads that either halt boiling or exceed maximum vapor flows in the overhead stream. PI and PID controllers do not operate well in such systems because they must be detuned to satisfy the constraint bounds. The authors pursue a model-predictive control strategy using a transfer function form of the energy balance to predict the energy produced by crystallization in the system. This allows prediction of the exothermic nucleation events that disturb the temperature of the system. After extensive tuning of the controller, the algorithm is shown to provide good temperature tracking in the crystallization of a weed killer-methanol system. However, the method does remain somewhat ad hoc as the controller is turned off during nucleation events.

The only batch crystallization study to attempt to describe CSD evolution in both crystal growth and dissolution regimes is the recent paper by Heffels and de Jong [62]. These authors formulate a discrete population balance with random growth rate fluctuation and postulate the existence of a nuclei population in which some nuclei are actively growing and some are not. This results in an effective nucleation rate. Unfortunately, little information is given as to how the parameters for the kinetic models are estimated and the resulting model simulations of an aqueous sugar system are unclear. However, the authors do present two interesting experimental results for an aqueous potash-alum system. They use sieve measurements to show that operating policies incorporating a dissolution phase into the batch profile can decrease the relative number of small crystals in the product CSD and improve crystal shape by breaking agglomerates.

A final note on batch control is provided by Bohlin and Rasmuson [10]. These authors use simulation to examine the industrial problem of implementing controlled cooling in systems where the kinetics are poorly determined and temperature control is not precise. By controlled cooling, the authors refer to curves with low initial cooling followed by strong cooling at the end of the crystallization. Their results indicate that, in general, cooling policies of this type are more vulnerable to errors in nucleation parameters or initial states, than are those of linear or natural cooling. These results point out the importance of characterizing the kinetics of a chemical system if performance is to be optimized successfully.

A general review of the crystallization control literature, for both continuous and batch systems, is provided by Rawlings et al. [124].

5.2 Calculation of Optimal Control Profiles

The optimal, open-loop control profiles are determined by solving a nonlinear program similar to that used for the optimal experimental design calculations described in Section 4.8.3.

Again, consider the initial conditions (e.g. initial concentration, seed mass, etc.) and duration of the experiment to be fixed and define the elements of the manipulated variable vector, $\boldsymbol{T}$, to be the vertices of a piecewise linear temperature profile. With constraints consisting of simple bounds on the elements of $\boldsymbol{T}$ and the corresponding cooling rates, a nonlinear path constraint on relative supersaturation, and final-time constraints on the yield, the optimal design problem may be restated as the following NLP:

$$\begin{aligned} \min_{\boldsymbol{T}} \quad & \Phi_c(\boldsymbol{T}, \mathbf{x}(t_f); \hat{\boldsymbol{\theta}}) \\ \text{subject to:} \quad & \text{Crystallizer Model} \\ & \boldsymbol{b}_l \leq \boldsymbol{A}\boldsymbol{T} \leq \boldsymbol{b}_u \\ & S(t) \leq S_{\max} \\ & \hat{C}(t_f) \leq \hat{C}_{\min} \end{aligned} \tag{NLP3}$$

where $\boldsymbol{A}$ is a constant linear constraint matrix similar to that described in Section 4.8.3 and $\hat{C}(t_f)$ is the final-time concentration in the liquid phase. $\Phi_c(\boldsymbol{T}, \boldsymbol{x}(t_f); \hat{\boldsymbol{\theta}})$ defines a function of the final-time model states that correlates with the desired performance of the system.

5.2.1 Control Objectives

This study is concerned with facilitating the filtration characteristics of the slurry. Unfortunately, it is not clear what form $\Phi_c(\boldsymbol{T}, \mathbf{x}(t_f); \hat{\boldsymbol{\theta}})$ should take to accomplish this goal. The flow of liquid through a cake of particles is a complex process that has been shown to depend on many factors including particle size distribution, particle shape, surface texture, and cake compressibility [145]. Clearly, these variables cannot be captured fully by a scalar function of the states of the CSD. However, under the assumption that other variables remain constant within a given system, particular CSD effects may be examined through simple functions of the model states. This study examines two such functions, the maximization of final-time seed size and the minimization of the ratio of the nucleated crystal mass to mass of grown seeds.

As noted in Section 5.1, Ajinkya and Ray [2] and Jones [80] manipulate temperature profiles to maximize the average size of the seed crystals. In the crystallizer model used in this study, this goal is equivalent to maximizing the integral of the growth rate over the course of a crystallization. The objective function to be minimized is

$$\Phi_c = -\int_0^{t_f} G(t)dt. \tag{5.1}$$

The positive form of this integral is the formal definition of the state ΔL that was introduced in Section 4.6.1 and therefore Equation 5.1 is hereafter referred to as the ΔL objective function.

A possible weakness of the ΔL objective is that it does not explicitly minimize the production of nuclei. Jones et al. [82] show that in the bimodal distributions of seeded crystallizers, the relationship between the mass fraction of the seeds and that of the nuclei may have a significant effect on the permeability of the filter cake. This leads Miller and Rawlings [94] to investigate the minimization of nuclei mass relative to that of the seeds. This metric is investigated in this study and has the objective function

$$\Phi_c = \left.\frac{m_N}{m_S}\right|_{t_f} \tag{5.2}$$

where m_N is the mass of the nucleated crystals and m_S is the mass of the grown seeds.

When the m_N/m_S objective is used, NLP3 has a trivial solution calling for zero yield. Thus, a yield constraint ($\hat{C}(t_f) \leq \hat{C}_{\min}$) must be specified if Equation 5.2 is to be an effective objective function.

5.2.2 Initial Conditions, Constraints and Profile Parameterization

The optimal profiles illustrated in the following sections were determined subject to a common set of initial conditions and constraints on the decision variables. As with previous experiments, the initial concentration for each profile was 0.11 grams photochemical/gram heptane. The crystallizer temperature was limited to a range between 18.5 and 24.0° Celsius. The temperature constraint assured that the optimal profiles would lie within the range of the solubility curve and that the assumption of temperature-independent kinetics would be preserved. Constraints also limited the optimal profiles to heating or cooling rates less than 0.1° Celsius/minute. However, because of the length of the experiments and spacing of the vertices (see below), the rate constraints were not active in any of the optimal profiles. The run length was 550 minutes and the piecewise-linear profile was divided into ten 55 minute segments, giving a decision variable vector $\boldsymbol{T}$ with 11 elements.

In each case, nonlinear constraints were used to limit the maximum level of relative supersaturation to 1.0 and to specify a final-time concentration of 0.056 grams/gram. The latter constraint corresponds to the lowest final-time concentration obtained during the model identification experiments and represents a yield of approximately 50%. The seed mass, taken to be 0.25 grams unless noted otherwise, equaled the highest level used in the identification experiments.

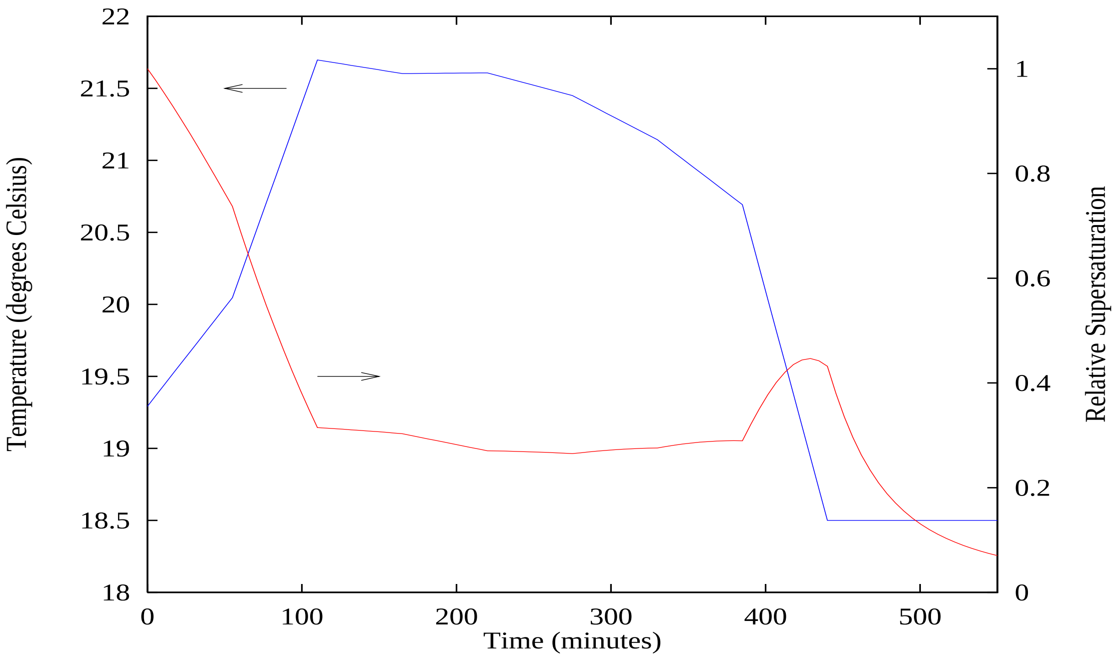

Figure 5.1: Simulated temperature and relative supersaturation profiles maximizing ΔL using the optimal model of Table 4.11.

5.3 Optimal Profiles for the Identified Model

Using the model parameters given in Table 4.11, optimal profiles were determined for both the ΔL and m_N/m_S objective functions. The optimal temperature and relative supersaturation profiles corresponding to these objectives are given in Figures 5.1 and 5.2, respectively.

The optimal profiles were determined by optimizing with respect to the elements of the temperature vector $\boldsymbol{T}$. However, because the mechanistic explanations of the optimal strategies are more easily expressed in terms of the relative supersaturation profiles, much of the discussion that follows focuses on supersaturation. Note the similarity between the temperature and supersaturation profiles determined for the two objective functions. This similarity is the result of the common need to suppress the nucleation process in order to

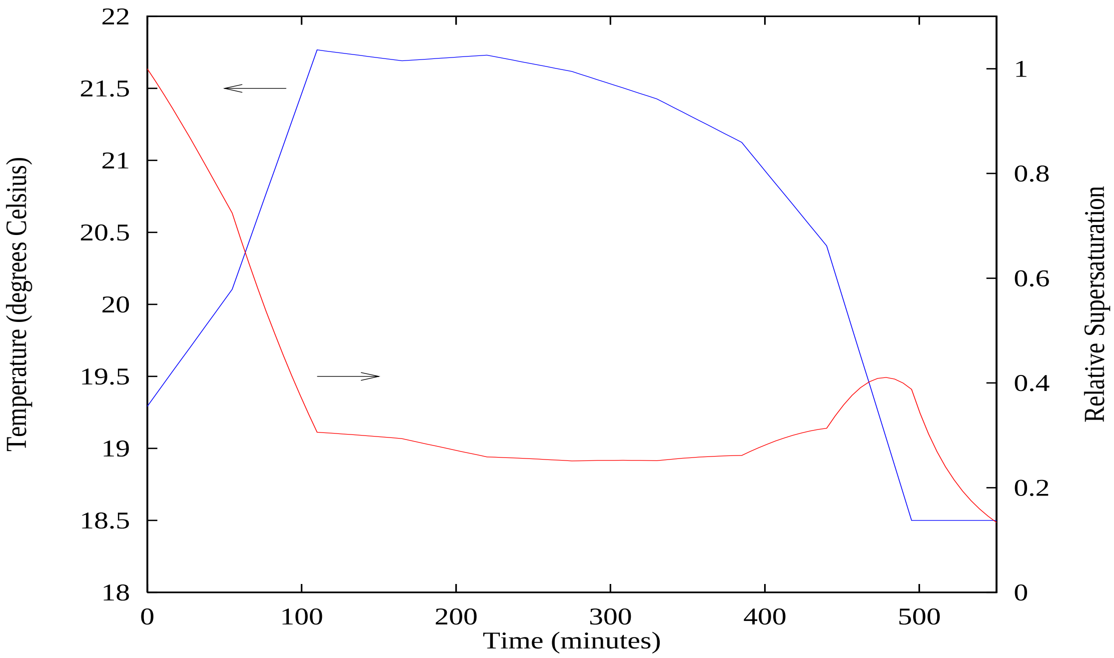

Figure 5.2: Simulated temperature and relative supersaturation profiles minimizing m_N/m_S using the optimal model of Table 4.11.

favor crystal growth. The optimal strategies each have three distinct parts; an initial period of high but declining driving force, an intermediate driving force plateau, and a final driving force peak.

The initial high supersaturation period is motivated by the dependence of the nucleation rate on the presence of large crystals. Since the initial seed mass is small and there are few large crystals early in the run, the optimal strategy capitalizes on the opportunity to allocate all of the supersaturation resource to crystal growth. As the crystals grow and the mass of large crystals increases the nucleation rate begins to rise, forcing the optimal policy to reduce the supersaturation in the system. The supersaturation is reduced because nucleation is higher order in supersaturation than is growth and is thereby favored by higher driving forces.

Once there is a significant population of large crystals, the supersaturation levels are dictated by trade-offs between the growth rate required to reach the yield constraint and the need to suppress nucleation. The optimal level is a function of $\hat{C}_{\min}$ and the difference in the supersaturation orders of growth and nucleation. The latter effect is investigated in Section 5.4.3. The supersaturation peaks at the end of the runs are used to drive the system to the yield constraint. The peak is positioned at the end of run because any nuclei generated so late in the experiment will have little time to accumulate significant mass.

Despite the similarities between the runs, note that the ΔL profile is more aggressive, maintaining a slightly higher driving force plateau for most of the run and producing a stronger driving force peak at the end. These differences are related to the lack of an explicit penalty on nucleation for the ΔL objective.

The seed mass and yield constraint used to determine the optimal profiles correspond to the seed mass and final-time concentration for the A-Optimal identification experiment. Table 5.1 presents the predicted final-time values of ΔL and m_N/m_S for the A-Optimal experiment and the optimized profiles. Relative to the A-Optimal profile, the optimal strategies each have higher final-time values of ΔL and lower final-time values of m_N/m_S. However, the improvements in the m_N/m_S ratio are more dramatic with the optimal profile predicting improvements over the A-Optimal and ΔL profiles of 32% and 11%, respectively. Relative to the m_N/m_S strategy, the profile optimizing $\Delta L(t_f)$ gives a surprisingly small increase in its objective function value. Since the max ΔL strategy appeared to provide no distinct benefits relative to the m_N/m_S strategy, it was not investigated further.

	$\Delta L(t_f)$ (μm)	$m_N(t_f)/m_S(t_f)$
A-Optimal	25.0	0.85
$\max \Delta L(t_f)$	27.0	0.65
$\min m_N/m_S$	26.9	0.58

Table 5.1: Comparison of final-time values of ΔL and m_N/m_S for control profiles and the A-Optimal identification experiment.

5.4 Parametric Analysis of m_N/m_S Optimal Profiles

It is instructive to examine the effects that optimization constraints and model parameters have on the optimal control strategy. While there are many factors to be considered, three of the most influential include the effects of the initial seed mass, the profile duration (t_f), and the magnitude of the difference in the supersaturation orders of growth and nucleation ($b - g$). The factors are discussed in the following sections.

All discussions relate to control strategies optimizing the final-time value of m_N/m_S. The nominal parameter set is taken to be the optimal disturbance model of Table 4.11. As with the previous section, control profiles are discussed in terms of their corresponding relative supersaturation trajectories.

5.4.1 Effect of Seed Load

Seed loading is an important element of industrial crystallization because it involves a trade-off between the solid-liquid separation performance of the process and the ultimate yield. In effect, whatever mass of product is sacrificed to seed a later crystallization amounts to a reduction in yield. The seed load used to determine the nominal optimal profile was 0.25 grams. This analysis compares the nominal profile and objective function value to similar strategies determined for seed loads of 0.15 and 0.35 grams.

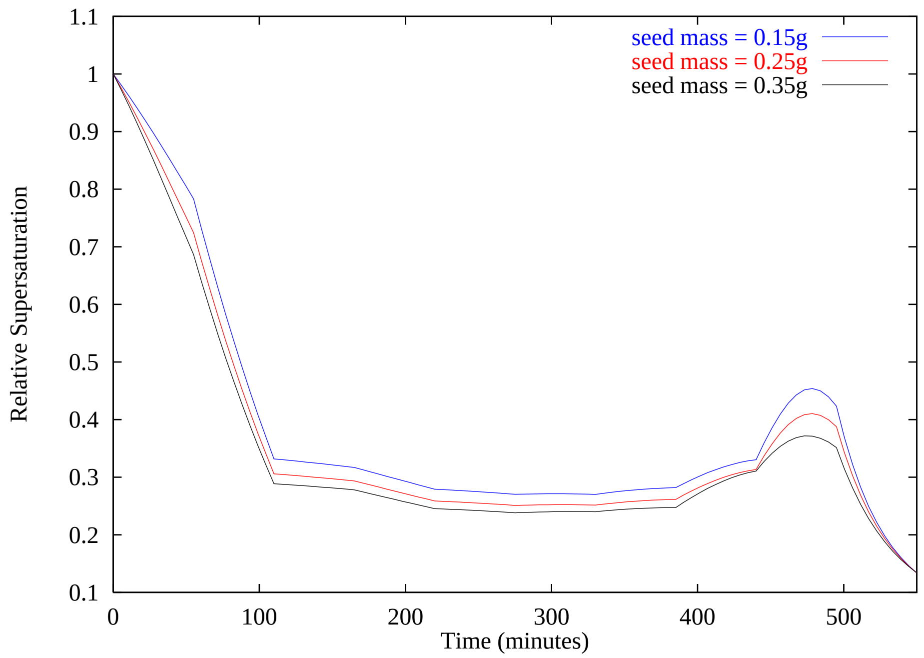

Figure 5.3: Optimal relative supersaturation profiles determined for seed loads of 0.15, 0.25, and 0.35 grams.

Seed Mass $\Rightarrow$	0.15 grams	**0.25 grams**	0.35 grams
$m_N(t_f)/m_S(t_f)$	0.98	**0.58**	0.38

Table 5.2: Effect of seed mass on the optimal value of the $\min m_N(t_f)/m_S(t_f)$ objective function with the nominal profile values in bold. The corresponding optimal relative supersaturation profiles are given in Figure 5.3.

The relative supersaturation profiles for the three seed loads are illustrated in Figure 5.3. The differences in the profiles are trivial and are driven entirely by the need to meet the yield constraint. Specifically, smaller seed loads require higher supersaturations to complete the transfer of mass into the crystalline phase in a limited time frame. However, despite the similarity of the optimal strategies, the differences in the final-time values of the objective function are dramatic. These values of m_N/m_S are compiled in Table 5.2.

Lowering the seed mass by 40% resulted in a 69% increase in the objective function.

A corresponding 40% increase in seed mass lowered the objective function by 34%. These results suggest that, if the m_N/m_S ratio is an important indicator of filtration qualities, higher seed loads are one way to achieve significant improvements. However, given the monotonicity of the effect (the assymptotic limit is to take the seeds as your product), the key to optimizing the seed load for a crystallization process is an accurate economic model that evaluates solid-liquid separation performance versus lost yield.

5.4.2 Effect of Profile Duration

Similar to seed load, the minimization of processing time is a variable of interest to industrial practitioners. It affects slurry separation properties, the operating costs of the crystallizer and related downstream processes, and affects batch scheduling strategies. The nominal strategy of Figure 5.2 was determined for a total crystallization time of 550 minutes. This duration was selected to allow comparisons with the previous identification experiments. The profile duration is analyzed by calculating similar operating profiles with final times that are 50 and 100% longer than the nominal duration.

The relative supersaturation profiles for the three profile durations are given in Figure 5.4. Clearly, the optimal strategies are extremely similar and are merely extended over the duration of the run. However, it is notable that as the value of t_f is increased the relative supersaturation levels decrease monotonically. This decrease is a response to the fact that nucleation is higher order in supersaturation than is growth. Thus, low supersaturations favor growth over nucleation. As t_f increases there is more time available to meet the yield constraint, allowing the optimizer to use lower the growth rate accordingly. As is shown in Table 5.3, the objective function benefits of going to progressively longer

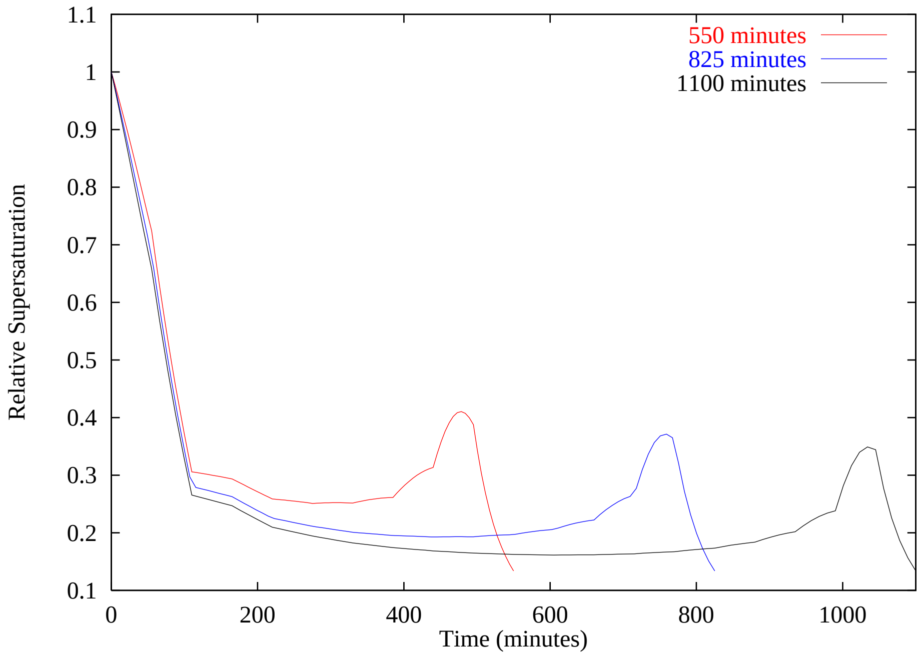

Figure 5.4: $m_N(t_f)/m_S(t_f)$-optimal relative supersaturation profiles determined for different values of t_f. All profiles determined with 55 minute piecewise linear temperature segments.

$t_f \Rightarrow$	**550 min**	825 min	1100 min
$m_N(t_f)/m_S(t_f)$	**0.58**	0.52	0.48

Table 5.3: Effect of t_f on the optimal value of the $\min m_N(t_f)/m_S(t_f)$ objective function with the nominal profile values in bold. The corresponding optimal relative supersaturation profiles are given in Figure 5.4.

times are moderate in comparison to seed load increases.

As with seed mass, simultaneous optimization of the run duration requires the specification of an economic objective function. Such a cost function would need to evaluate the solid-liquid separation improvements relative to increased operating costs and scheduling concerns.

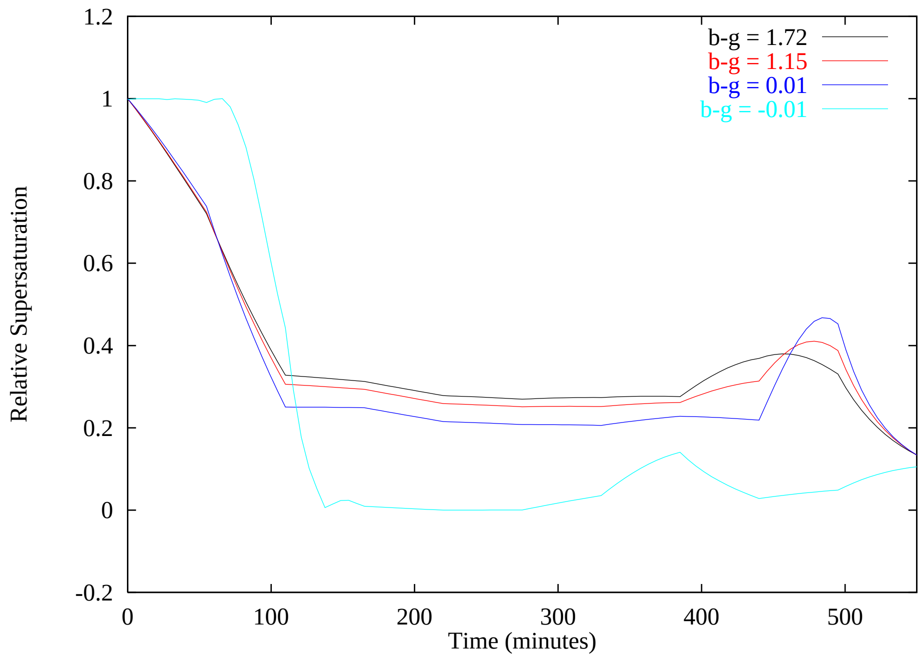

Figure 5.5: Optimal m_N/m_S relative supersaturation profiles determined for different values of $(b-g)$.

5.4.3 Supersaturation Orders of Nucleation and Growth

Unlike seed mass and profile duration, the parameters b and g are defined in the identified model. Though they may not be adjusted to improve the optimization, the difference in their optimal values dictates both the shape of the optimal control profile and the degree of objective function improvement that may be expected from its implementation. To assess the effects of the difference quantity $(b-g)$, control strategies were determined for a large positive value of $(b-g)$ (150% of the nominal value), a small positive value of 0.01, and a small negative value of -0.01. These profiles and their optimal objective function values are presented in Figure 5.5 and Table 5.4, respectively.

The striking feature of Figure 5.5 is the change in the form of the optimal relative supersaturation profile as the value of $(b-g)$ becomes negative. In this state, crystal growth is favored by high supersaturations and the optimal policy takes advantage by driving the system with the maximum allowable value of relative supersaturation. High driving force leads to a rate of mass transfer into the solid phase that nearly satisfies the yield constraint within 120 minutes of the start of the run. For the remainder of the run, the optimal profile minimizes the driving force to prevent nucleation. The slight driving force peaks at the end of the run are used to meet the yield constraint exactly. However, it is not clear that this portion of the optimal profile is well-determined.

Unlike previous optimal control profiles, NLP3 fails to converge for the negative $(b-g)$ optimal policy. The lack of convergence may occur because the optimal policy is attempting to maintain the driving force on the nonlinear constraints (supersaturation levels of 0 and 1) with a rough temperature discretization. A second factor complicating the identification of the optimal profile is the insensitivity of the CSD to relative supersaturation. Note in Table 5.4 that despite dramatically different temperature profiles, the objective functions determined for the 0.01 and -0.01 profiles are essentially identical. The similarity of the objective functions may be explained through a brief analysis of the zero-size boundary condition for the PBE. The boundary condition has the form

$$f(0,t) = \frac{B^\circ}{G|_{L=0}} = \frac{k_b}{k_g} S^{b-g} \mu_3^j$$

where S is the relative supersaturation. Thus, as $(b-g)$ approaches zero the boundary

$b-g \Rightarrow$	-0.01	0.01	**1.15**	1.72
$m_N(t_f)/m_S(t_f)$	0.91	0.90	**0.58**	0.46

Table 5.4: $m_N(t_f)/m_S(t_f)$ evaluated for the $(b-g)$ profiles of Figure 5.5. The values corresponding to the nominal profile are in bold.

condition ceases to be a function of relative supersaturation, meaning that, given the same seeds and yield constraints, when $b - g \approx 0$ all temperature profiles produce the same product CSD. The implication is that the expected benefit of employing optimal temperature profiles is a direct function of the values of b and g determined for a particular system. The expected benefits are illustrated by the objective function values of Table 5.4. As $(b - g)$ increases, progressively greater benefits are conferred through lower optimal values of $m_N(t_f)/m_S(t_f)$.

Finally, note that for the positive $(b - g)$ profiles the driving force level applied during the plateau portion of the supersaturation trajectory increases as $(b - g)$ increases. This response is related to the mass order of nucleation, j. With $j \approx 2$, the nucleation penalty for high supersaturations late in the run (when crystal mass is greatest) is extreme and it grows more so as $(b - g)$ increases. Therefore, the optimal profiles trade-off increasing amounts of nucleation throughout the run in order to minimize the height of the expensive supersaturation peak at the end.

5.5 Robustness of the Optimal Strategy

The analysis of the optimal control policy is completed by assessing its sensitivity to the model parameters. A simple way to examine this sensitivity is to compare the control profile determined for the optimal parameters to those determined for the alternate models introduced in Table 4.8. Recall that these models are comprised of optimal parameters determined for data sets containing 25, 50, and 100% of the concentration data from the identification experiments that were biased by the caking of crystals on the densitometer filter. Figure 5.6 illustrates that despite the differences in their modal parameter sets, the

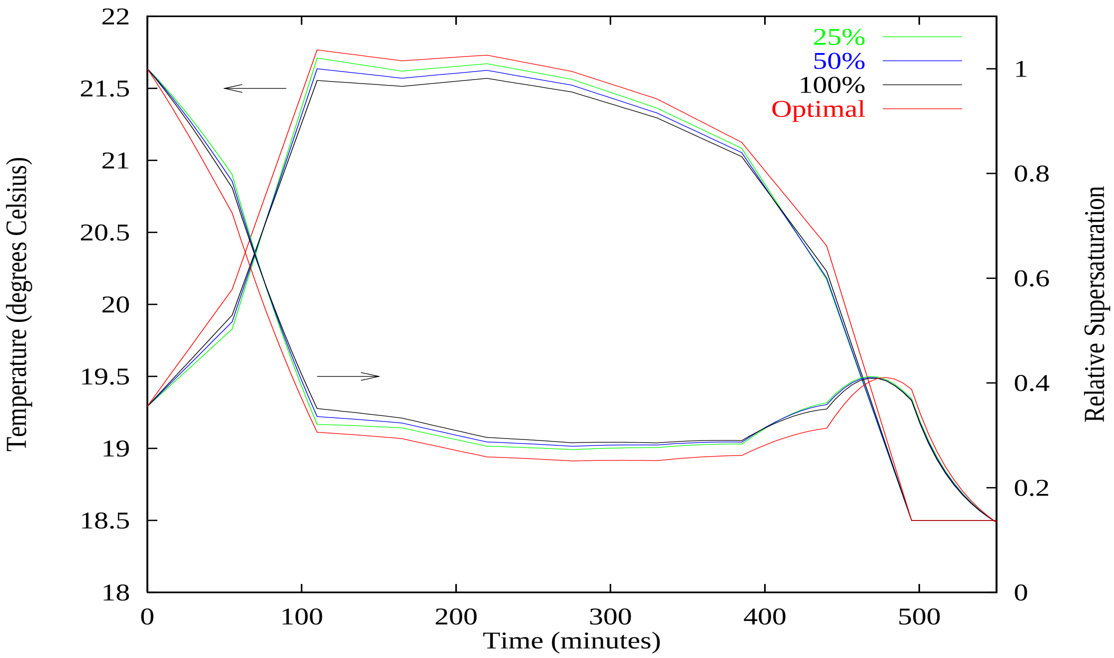

Figure 5.6: Simulated temperature and relative supersaturation profiles determined for the models of Table 4.8 and the optimal disturbance model of Table 4.11.

alternate models produce similar optimal profiles. This result indicates that the optimal strategy is relatively insensitive to the model decision discussed in Section 4.10.

A more formal accounting of the robustness of the control policy may be performed by modeling the optimization of the control profile, $\boldsymbol{u}$, as a nonlinear mapping from the parameter space into the control space. Let $\boldsymbol{\delta u}$ be the difference between the optimal profile determined for parameter set $\boldsymbol{\theta}$ and that determined for the nominal set $\hat{\boldsymbol{\theta}}$:

$$\boldsymbol{\delta u} = \mathcal{N}(\boldsymbol{\theta}) - \mathcal{N}(\hat{\boldsymbol{\theta}}).$$

Linearizing $\mathcal{N}$ about the nominal parameter set allows an approximate statement of the optimal control problem:

$$\boldsymbol{u} \approx \boldsymbol{L}\boldsymbol{\theta}$$

$$L_{ik} = \left.\frac{\partial u_i}{\partial \theta_k}\right|_{\hat{\boldsymbol{\theta}}}.$$

From an analysis of the linearized control problem, it is possible to determine the perturbation direction from $\hat{\boldsymbol{\theta}}$ that is amplified to the greatest extent by the matrix $\boldsymbol{L}$. This amplification produces the $\boldsymbol{u}$ that is most displaced (in a two-norm sense) from its nominal value. The exact displacement of $\boldsymbol{u}$ is determined by varying the nominal parameter set along the most highly amplified direction and resolving the nonlinear control problem.

According to the linear approximation,

$$\boldsymbol{\delta u} = \boldsymbol{L}(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}) = \boldsymbol{L}\boldsymbol{\delta\theta}.$$

Performing a singular value decomposition of $\boldsymbol{L}$ gives,

$$\boldsymbol{\delta u} = \boldsymbol{Q}_1 \boldsymbol{\Sigma}\ \boldsymbol{Q}_2^T \boldsymbol{\delta\theta}.$$

The largest singular value of $\boldsymbol{L}$ (the first element on the diagonal of $\boldsymbol{\Sigma}$) defines the magnitude of the maximum amplification of the matrix $\boldsymbol{L}$. This amplification is applied to the quantity $\boldsymbol{\delta\theta}_{\mathrm{worst}}^T\boldsymbol{\delta\theta}$, where $\boldsymbol{\delta\theta}_{\mathrm{worst}}$ is the first column of the right singular matrix $\boldsymbol{Q}_2$. Clearly, the $\boldsymbol{\delta\theta}$ vector maximizing $\boldsymbol{\delta\theta}_{\mathrm{worst}}^T\boldsymbol{\delta\theta}$ is a multiple of $\boldsymbol{\delta\theta}_{\mathrm{worst}}$ itself. Therefore, the $\boldsymbol{\delta\theta}$ that maximizes $\boldsymbol{\delta u}^T\boldsymbol{\delta u}$ is a multiple of the first column of the right singular matrix, $\boldsymbol{Q}_2$. Similarly, the second worst perturbation vector is the second column of $\boldsymbol{Q}_2$. The relative severities of the worst and second worst perturbations are dictated by the relative magnitudes of the first and second singular values.

For the nominal parameter set, the worst-case perturbation vector is

$$\boldsymbol{\delta\theta}_{\text{worst}} = [0.483 \;\; -0.566 \;\; 0.112 \;\; -0.130 \;\; -0.645 \;\; -0.016 \;\; -0.021]^T .$$

Note that $\boldsymbol{\delta\theta}_{\text{worst}}$ has only seven elements because the disturbance parameter k_f was only nonzero for the biased identification experiments. Figure 5.7 gives the results of solving the optimal control problem after perturbing the nominal parameters by $\pm 0.1(\boldsymbol{\delta\theta}_{\text{worst}})$. The perturbations produce noticeable changes in the optimal profile, however, the basic trends of the nominal profile are preserved. The corresponding relative supersaturation profiles are similar to those shown in previous sections, with the negative and positive perturbation profiles initiating the late driving force peak earlier and later, respectively, than the nominal strategy. To demonstrate the decrease in the severity of the $\boldsymbol{Q}_2$ directions, the results of $\pm 10\%$ perturbations along the second worst direction are also plotted in Figure 5.7. The singular value amplifying the worst direction is 10.7, while the second singular value is 2.0.

The worst-case deviations depicted in Figure 5.7 understate the robustness of the optimal control strategy. This is because the $\pm 0.1(\boldsymbol{\delta\theta}_{\text{worst}})$ deviations are outside of the 95% confidence hyperellipsoid predicted by Equation 4.25. For example, note that the +10% deviation along the g axis (the second element of $\boldsymbol{\delta\theta}_{\text{worst}}$) lowers the nominal estimate of g by 0.0566. This is well-outside of the ± 0.01 conservative 95% confidence interval for g. The worst-case deviations within the ellipsoidal HPD region may be derived from Equation 4.25, giving

$$\boldsymbol{\theta} = \hat{\boldsymbol{\theta}} \pm \lambda \boldsymbol{\delta\theta}_{\text{worst}} \tag{5.3a}$$

$$\lambda = \left(\frac{\chi_p^2(\alpha)}{\boldsymbol{\delta\theta}_{\text{worst}}^T \boldsymbol{H}_{\hat{\theta}} \boldsymbol{\delta\theta}_{\text{worst}}} \right)^{1/2} \tag{5.3b}$$

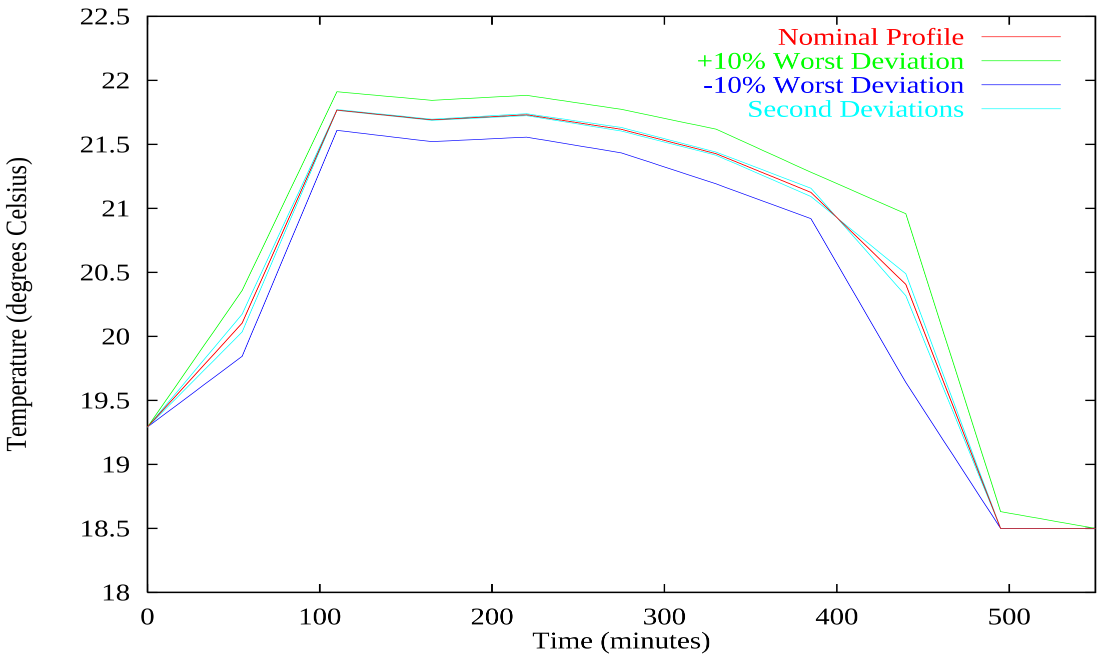

Figure 5.7: Optimal profiles determined for the nominal parameter set and $\pm10\%$ parameter deviations along $\boldsymbol{\delta\theta}_{\text{worst}}$. Profile change due to $\pm10\%$ deviations along the second worst direction are also plotted.

where λ is the magnitude selecting the intersection of $\boldsymbol{\delta\theta}_{\text{worst}}$ and the HPD region boundary. For the nominal linear 95% HPD region, $\lambda = 0.0022$. The control profiles corresponding to positive and negative perturbations of this magnitude are indistinguishable from the nominal temperature trajectory, which suggests that the optimal control profile is robust with regard to the uncertainty of the nominal parameter estimates.

5.6 Implementation of Optimal Control Policies

Two optimal control strategies minimizing the $m_N(t_f)/m_S(t_f)$ objective function were implemented on the crystallization apparatus described in Chapter 2. Each control policy was

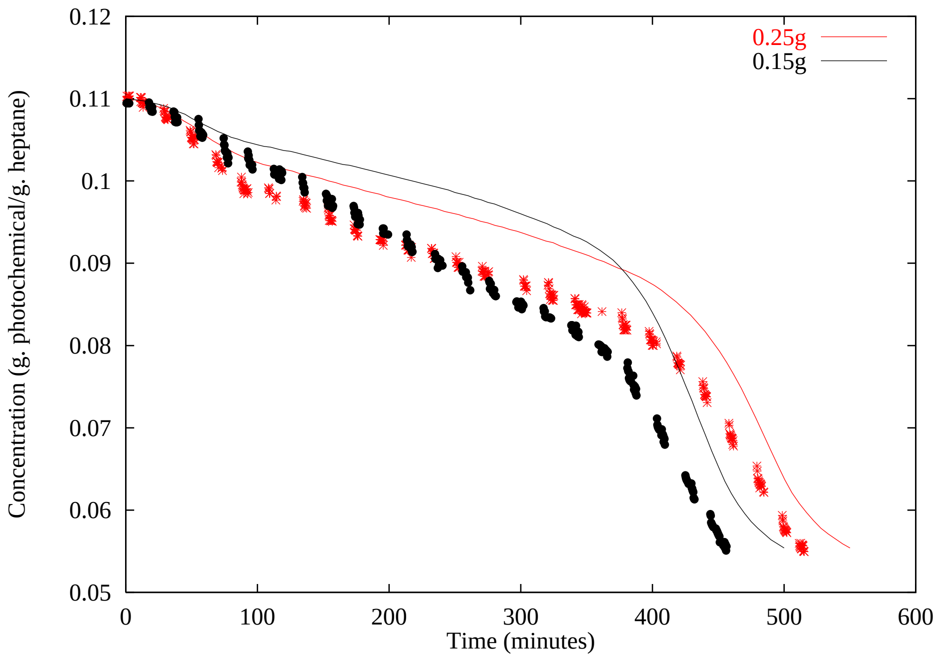

Figure 5.8: Concentration data and model predictions for the optimal control policies minimizing $m_N(t_f)/m_S(t_f)$.

determined subject to the parameters of the 25% concentration model given in Table 4.8.[1] The first policy was a 550 minute run with a 0.25 gram seed load. This profile is given in Figure 5.6 and is similar to that later determined for the optimal disturbance model. The second optimal profile was determined for a 500 minute run with a 0.15 gram seed load.

The model fits to the concentration data for the optimal control experiments are given in Figure 5.8. Clearly, there is significant model mismatch in these experiments. Because the reason for the large residuals was not immediately clear, the experimental system checked by conducting an additional replicate of the type 4 identification experiment (see Section 4.5). Figure 5.9 shows the concentration data collected from this replicate experiment, denoted 4d, plotted alongside data from earlier replicates 4a, 4b, and 4c.

[1]The optimal disturbance model was identified later during a reexamination of the bias problem. The 25% model was chosen over the 50 and 100% models because because it provided a roughly equal weighting of the biased and unbiased identification experiments (see discussion in Section 4.10).

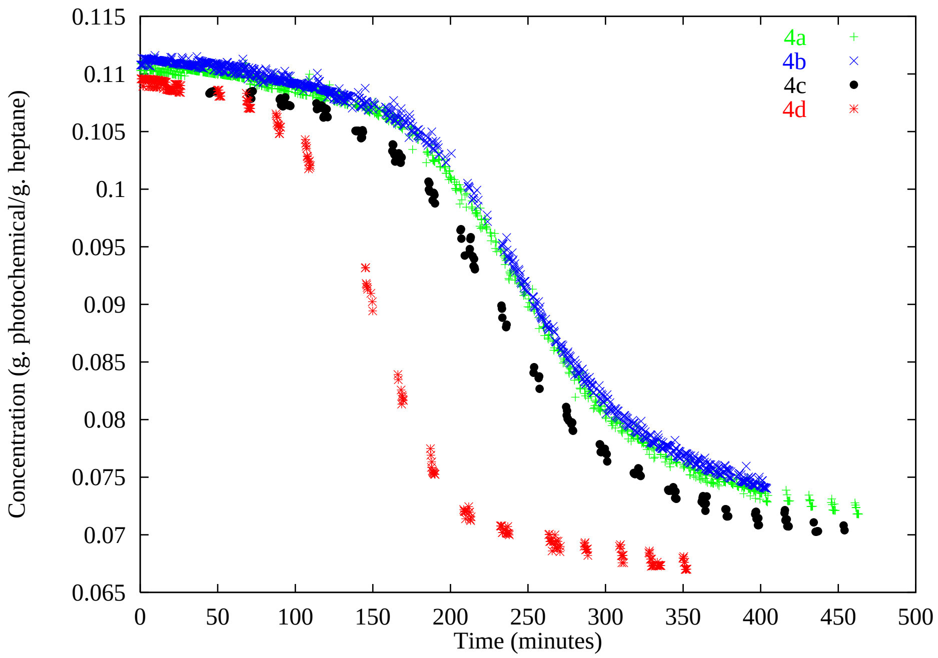

Figure 5.9: Comparison of concentration data for replicate experiments 4a-d. Replicate 4d was conducted after the optimal control experiments.

Figure 5.9 indicates that there was a significant change in the system kinetics between the time of the identification experiments and the initiation of the control study. This change was responsible for the model mismatch observed in Figure 5.8. The only known alteration in the experimental conditions over this time was a change in the container from which the photochemical raw material was drawn, suggesting that there was a difference in the raw material stocks in the two containers.

As noted in Section 4.8.4, one motivation for halting the identification study was the fear of exhausting the limited supply of raw material stock. Recall that the stock was collected at the production plant by Kodak. After delivery to the laboratory, nearly all of the raw photochemical was transferred from the original container (a heavy plastic bag) into a sealable 5 gallon plastic container. In order to assure a uniform stock of raw material, the photochemical supply in the plastic container was stirred regularly. This stirred stock was the material used and nearly exhausted by the identification study. After the completion of the identification study, the only remaining photochemical stock was that which had been left in the original container. This material was combined with the remaining stirred stock to provide sufficient material for the control experiments. The ensuing change in the system kinetics was assumed to be due to this change of material.

Given the shift in the system kinetics and the near exhaustion of the remaining supply of raw material, the control study was halted after the source of the model mismatch was identified. However, despite the change in the system kinetics, the filtration results for the two optimal control profiles were still of interest. A comparison between the filtration qualities of the optimized slurries and those from the identification experiments is presented in the following section.

5.7 Comparative Summary of Filtration Data

As noted in Chapter 2, quantitative measurements of the qualities of the crystal size distribution are not available for the photochemical system. Thus, it was impossible to measure experimentally the $m_N(t_f)/m_S(t_f)$ objective function or even to evaluate a related function such as the weight-mean size. For this reason, this study attempted to assess the effectiveness of the optimal control strategy through the determination of the average specific cake resistance, denoted $\overline{\alpha}_c$. This parameter is determined according to Equation 2.13 and is preferred over other metrics (e.g. filtration time) because it is independent of solids concentration and thereby allows the comparison of slurries with different yields.

run type	# samples	$\overline{\alpha_c}$ (cm/g)	% solids	time (sec)
1	1	1.97×10^9	5.8	54.0
2	2	2.51×10^9	7.3	99.0
3	2	1.66×10^9	5.8	62.0
4	2	2.72×10^9	5.8	75.1
5	1	2.56×10^9	8.4	81.4
A-Opt	1	2.49×10^9	8.7	92.9
$\Phi - .15$g	1	1.44×10^9	8.9	58.7
$\Phi - .25$g	1	1.14×10^9	8.9	51.8

A disadvantage of this type of measurement is that the results are more qualitative because the experimentally determined measure of improvement is not the same, or even clearly related to, the optimized objective function. Thus, the cake resistance can only determine conclusively whether or not slurries crystallized via the optimized policy filter better than do those crystallized with different policies. However, such a qualified conclusion is still significant when it comes from a direct measure of the quality of interest and is compared to a reasonable range of alternatives. In this study, the average cake resistances determined for the two optimized profiles are compared to resistances determined for the model identification experiments.

Table 5.7 presents a summary of the filtration data collected for both the identification experiments and the control experiments. Where replicates were available (i.e. runs 2, 3, and 4), the values in Table 5.7 are averages. Replicate experiments exhibited $\overline{\alpha}_c$ variations of 10-20%. This variation emphasizes that filtration measurements tend to be noisy and difficult to replicate. In this study, identifiable sources of this noise included the variation in the CSD's of replicate runs (shown repeatedly in the residual plots of Chapter 4) and mild differences in the constant pressure drop across the filter cake (noted in Section 2.6.2).

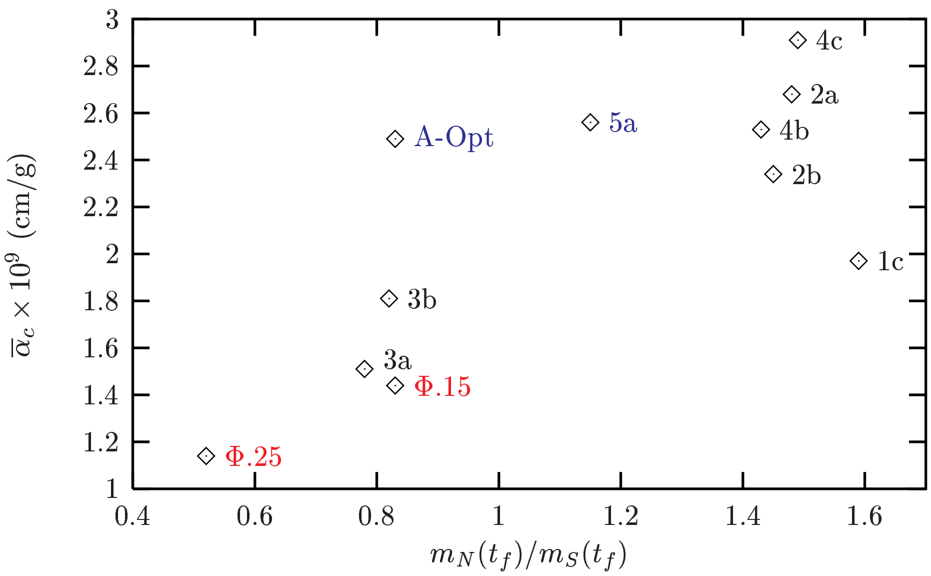

Figure 5.10: Correlation between average specific cake resistance and predicted values of $m_N(t_f)/m_S(t_f)$. Runs optimized for control are in red while those optimized for estimation are in blue.

Table 5.7 shows that the lowest cake resistance values recorded in this study corresponded to the two optimal profiles. The $\Phi - .25$ experiment has a cake resistance 25% lower than any other experiment in the study. Also, despite having higher solids densities, the optimized slurries filtered faster than did all but one of the previous experiments. These points confirm that the optimization process does improve the filtration qualities of the resulting slurries. Further evidence is provided by plotting $\overline{\alpha}_c$ versus the $m_N(t_f)/m_S(t_f)$ objective function for all of the filtered slurries. This plot is given in Figure 5.10. The optimal control profiles are labeled in red while the optimal identification experiments are in blue.

The data in Figure 5.10 are noisy and care should be taken in drawing specific conclusions. However, the data do suggest a correlation between $m_N(t_f)/m_S(t_f)$ and cake resistance. This correlation may be summarized by calculating the correlation coefficient

	m_N/m_S	c_v	c_{vw}	$\overline{L}$	$\overline{L}_w$	ΔL
correlation coefficient	0.73	-0.85	0.84	0.62	-0.91	0.35

Table 5.5: Correlation coefficients expressing the linear relationships between several predicted scalar measures of the final-time CSD and experimentally determined values of average specific cake resistance, $\overline{\alpha}_c$.

for the data in Figure 5.10. The correlation coefficient is a scaled measure of the variance of a sample about a least squares line. Therefore, the correlation coefficient merely indicates a degree of linear correlation between two variables and does not guarantee a causal relationship. However, correlation coefficients may indicate some general trends to suggest objective function choices for investigation in future studies. The correlation coefficient for $\overline{\alpha_c}$ versus $m_N(t_f)/m_S(t_f)$ is given in Table 5.7. Table 5.7 also presents correlation coefficients describing relationships between $\overline{\alpha_c}$ and predicted final-time values of the population coefficient of variation (c_v), weight coefficient of variation (c_{vw}), population mean size ($\overline{L}$), weight-mean size ($\overline{L}_w$) and ΔL.[2]

The results shown in Table 5.7 demonstrate the expected correlations for the crystal mass-based measures. Positive correlations are indicated between cake resistance and final-time values of m_N/m_S and c_{vw}. The positive c_{vw} correlation is consistent with the results of Jones et al. [82] that show monotonic improvement in cake permeability as c_{vw} decreases. A strong negative correlation with the final-time value of $\overline{L}_w$ indicates that the filtration is improved as the weight-mean size of the crystals increases. These results are countered by nonintuitive correlations for the population or number-based measures of c_v and $\overline{L}$. The strong negative correlation between $\overline{\alpha}_c$ and c_v suggests that increased size variation in the population aids filtration, implying that, at least in seeded batch crystallizers, reduction of

[2] $c_v = \sqrt{\mu_2\mu_0/\mu_1^2 - 1}$, $c_{vw} = \sqrt{\mu_5\mu_3/\mu_4^2 - 1}$, $\overline{L} = \mu_1/\mu_0$, and $\overline{L}_w = \mu_4/\mu_3$

c_v may actually harm filtration performance. A positive correlation with $\overline{L}$ is also counterintuitive, indicating that increased average size slows filtration. The counterintuitive nature of the population-based correlations may imply that population-based measures are inappropriate for predicting filtration characteristics.

The data in Table 5.7 and Figure 5.10 show that the filtration of the photochemical-heptane slurries have been improved by cooling policies minimizing the $m_N(t_f)/m_S(t_f)$ objective. Ideally, this result would be supported by a greater array of filtration replicates but this was not possible in this study due to the exhaustion of the raw material stock.

5.8 Summary

This chapter used the identified kinetic model of Chapter 4 to determine optimal cooling policies for the photochemical-heptane system. The optimal policies were intended to improve the filtration characteristics of the final-time slurries and were determined subject to constraints on cooling rate, driving force, and final-time yield. Objective functions maximizing final-time seed size and minimizing of the mass of nucleated crystal relative to the mass of grown seed crystals, $m_N(t_f)/m_S(t_f)$, were examined. The effects of seed load, run duration, and the supersaturation orders of nucleation and growth on the (m_N/m_S)-optimal policy were analyzed. The improvement of the objective function was shown to be monotonic with respect to seed load and run duration. The sensitivity of the m_N/m_S-optimal policy to uncertainty in the kinetic parameter estimates was also assessed.

Two optimal control experiments employing temperature profiles minimizing the m_N/m_S objective function were performed and shown to improve filtration relative to previous experiments. Filtration performance was assessed using the average specific resistance of the filter cake. The optimal control experiment using the higher seed load obtained a cake resistance 25% lower than any other experiment in the study. A plot of cake resistance versus $m_N(t_f)/m_S(t_f)$ for all of the filtration experiments in the study showed a rough correlation suggesting that lower values of $m_N(t_f)/m_S(t_f)$ produced lower flow resistance in the filter cake. Correlation coefficients were determined to relate experimentally determined values of average specific cake resistances to predicted final-time values of several scalar measures of the CSD (coefficient of variation, weight mean coefficient of variation, mean size, weight mean size, m_N/m_S, and seed size). These correlations suggested that mass-based objectives such as the minimization of m_N/m_S, minimization of weight-mean coefficient of variation or maximization of weight-mean size would be effective objectives for the minimization of flow resistance in the filter cake.

Chapter 6

Conclusions

As noted in the introduction, crystallization is an important solid-liquid separation technique used in the chemical process industries. Due to low operating costs and flexibility, batch crystallization processes are especially important for manufacturers of pharmaceuticals and photochemicals. These producers use batch crystallizers to separate product chemicals from complex reaction mixtures. Such crystallizations are often plagued by high impurity levels, needle-like crystal habits, and poorly characterized solubility curves. By contrast, most of the academic research in the field of crystallization has focused on describing the behavior of model systems such as inorganic salts. Such studies have established the utility of a population balance modeling approach for describing crystal size distribution in batch crystallization systems. However, these methods required testing on chemicals systems more similar to those found in industry. This study has applied a fundamental, population balance approach to identify and control the crystallization kinetics in an industrial chemical system. The system chosen was a photochemical-heptane system that is currently in production at the Rochester, New York facilities of the Eastman Kodak Company.

The approach used in this work involved the specification of a general model to describe the crystal size distribution in a laboratory-scale seeded batch cooling crystallizer. The general model consisted of a population balance model describing the size distribution of the crystals and mass and energy balances to allow the determination of the relative supersaturation driving force. Empirical, power law models were used to relate the kinetic processes of crystal growth and nucleation to relative supersaturation. Using on-line measurements of liquid-concentration and slurry turbidity (transmittance), the parameters in the kinetic models were estimated using a Bayesian inference technique. The identified model was then used to determine optimal open-loop control policies with the goal of improving the filtration properties of the final-time slurries. A measure of the flow resistance of the filter cake was used to characterize the quality of the filtrations. The following sections review the conclusions of the model identification and control studies in detail.

6.1 Model Identification

The model identification study showed that a fundamental modeling approach may be successfully applied to a problematic crystallization system. Kinetic models and parameters were identified to allow the general population balance model to predict accurately the concentration and transmittance measurements recorded for 11 seeded crystallization experiments. The kinetic processes of crystal growth and secondary nucleation were related to the relative supersaturation driving force by temperature-independent empirical power-law expressions.

Conventional rate constants and supersaturation exponents were estimated for both models. Two additional parameters were estimated for the nucleation process. These parameters described the dependence of the secondary nucleation process on the total mass of crystals above a particular size. The order of the mass term and the minimum size were estimated from the data. The size-dependent nucleation model was a variation of a model proposed by Ottens et al. [104]. According to Ottens et al., the minimum size corresponds to a minimum mass that must be obtained before crystal collisions may produce secondary nuclei.

Two additional parameters were also estimated from the data. The first of these parameters described the increase in the length-to-width ratios of the photochemical seeds as they underwent growth. This habit change was verified experimentally using optical microscopy of crystals collected at different times during a crystallization. The second parameter was used to model a disturbance that affected 7 of the 11 identification experiments. The disturbance was related to deposition of crystals on the filter that prevented crystals from entering the sample stream used for the concentration measurement.

The uncertainty in the estimated parameter set was evaluated using linear statistics. Linear 95% confidence intervals were determined by calculating the dimensions of a hyperbox circumscribing the hyperellipsoidal region of highest posterior density for the parameters. These confidence intervals provide a conservative measure of the uncertainty in the parameters. Tests on the hyperellipsoidal region demonstrated that it gave a reasonable representation of the nonlinear posterior density about the modal parameter set. It was subsequently shown for two parameter pairs that the conservative intervals defined by the hyperbox did enclose the regions of appreciable probability.

The linearity assumption was also used to design an optimal experiment to reduce the average variance of the optimal parameters. The optimal design provided a temperature profile that optimized the information in the data set by using strong temperature gradients and charting regions of the operating space that were not studied in previous experiments.

6.2 Optimal Control

Having identified the optimal kinetic parameters, the identified batch crystallizer model was used to determine optimal open-loop control profiles to improve the filtration of the final-time slurries. The focus was on the minimization of the final-time mass of nucleated crystals relative to the mass of grown seeds (m_N/m_S). The optimal profile was subject to constraints on the supersaturation range and the final-time yield.

The optimal profile was shown to reduce the predicted final-time value of m_N/m_S by 32% as compared to the optimal experimental design run, which had the same seed mass, profile duration, and yield. The sensitivity of the optimal profile to changes in seed mass and profile duration was assessed. It was shown that increases in seed mass or profile duration lead to further reductions in the m_N/m_S objective function. The most dramatic improvements corresponded to increased seed loads. The importance of the relative values of the supersaturation orders of the growth and nucleation processes was also demonstrated. The larger the absolute value of $(b - g)$, the greater the potential benefits of implementing optimal temperature profiles.

The sensitivity of the optimal control profile to the uncertainty in the parameters was assessed. The optimal profile was shown to be robust to the worst-case disturbance within the hyperellipsoidal confidence region. The worst-case disturbance was determined by modeling the nonlinear control problem as a linear process and performing a singular value decomposition on the sensitivity matrix that maps the parameters into the control trajectory.

The efficacy of m_N/m_S objective function for predicting improvements in filtration was tested by implementing two optimal control experiments on the batch crystallizer apparatus. Constant pressure filtration experiments conducted on the resulting slurries showed cake resistance improvements relative to 9 similar filtrations conducted for the identification experiments. The average specific cake resistance measured for the optimized experiment with the larger seed mass was shown to be 25% lower than any other resistance value recorded during the study.

The filtration results of the optimal control study were tempered somewhat by the failure of the model to describe the concentration profiles of the optimal control experiments. This failure was associated with a shift in the kinetics of the system that corresponded to the use of a different raw material stock. The new stock was of the same industrial batch but was stored in different container than the material used in the model identification experiments. The shift in the kinetics may have corresponded to different impurity levels in the two stocks.

6.3 Suggestions for Future Work

This project established that a fundamental modeling approach may be used to improve the filtration characteristics of a relevant industrial crystallization system. However, the improvements garnered in this study were incremental. One avenue for attaining more dramatic improvements is to allow the crystals to agglomerate. Though no crystal agglomeration was apparent in this study, one experiment conducted by the Eastman Kodak Company did form large agglomerates. When filtered, the agglomerates gave a cake resistance two orders of magnitude lower than those measured in this study [44]. The agglomerates were formed in a batch experiment with low agitation and high relative supersaturation.

While agglomeration is cumbersome to model via a formal population balance approach, Ramkrishna et al. [119] and Gooch and Hounslow [57] show that Monte Carlo simulation provides a convenient treatment of the phenomenon. It remains to be seen whether or not Monte Carlo simulations are computationally too intensive to allow the identification of agglomeration kernels from experimental data. Additional concerns for modeling an agglomerative system include the possible inadequacy of the ideal mixing assumption due to low agitation rates and the replicability of a process that may require the generation of primary nuclei [97]. Despite these challenges, the filtration improvements observed at Kodak merit further investigation of agglomeration effects.

A second area for potential study is the simultaneous use of solution temperature and fines dissolution in optimal control of batch crystallizers. As noted by Rohani et al. [130], fines dissolution may be used to improve mean crystal size but the yield loss may be substantial. The final-time constraints of the optimal control scheme developed by Miller and Rawlings [94] would allow for the incorporation of fines dissolution without the subsequent loss of final-time yield. The incorporation of fines dissolution would also allow for a measure of feedback control not possible with temperature alone. For instance, it would be possible to control final-time CSD in systems in which the supersaturation orders of growth and nucleation were approximately equal (recall the discussion of Section 5.4.3). Such a feedback scheme could then be incorporated with on-line state estimation to control disturbances from off-setpoint initial conditions or mild shifts in the kinetics of nucleation and growth. Bohlin and Rasmuson [10] have shown that such disturbances may significantly degrade the performance of optimal open-loop control profiles.

Lastly, there needs to be more work done on the identification of size distribution based objective functions for optimizing slurry filtration. This topic was investigated in this study and has also been treated by Jones et al. [82] and Jones and Mydlarz [83]. However, none of these studies have compiled a sufficient amount of filtration data to make strong comparisons between different objectives. This gap in the literature is understandable given that filtration provides the investigator with only a single data point per crystallized slurry. Therefore, filtration studies require a great deal of labor to produce a limited amount of relatively noisy data. Nonetheless, for a model chemical system that may be easily replicated and relatively easily identified, perhaps the KNO_3-H_2O system used by Miller [95], it may be possible to focus the study on controlling CSD for improved filtration. Because filtration is influenced by particle shape, the results of the initial filtration study could be tested in the optimization of a second system with a different crystal habit.

As a final note, parameter estimation in crystallization is hampered by the dearth of sensor technologies capable of measuring the states of the solid phase particles. Microscopy is one measurement that may offer some solutions to this dilemma. The determination of CSD from photomicroscopy is, in general, a tedious and difficult process. However, as shown in this study, microscopy may provide a valuable measurement of crystal shape to augment more traditional measurements of concentration and transmittance. On-line implementation of video microscopy is capable of generating real-time information about crystal shape, size and quality. On-line video microscopy may be particularly helpful for agglomerative systems because agglomeration allows increases in particle size without the transfer of solute from the liquid phase. Such increases would not be observable in a concentration measurement. Given these advantages, the addition of on-line microscopy to existing measurements of transmittance and concentration is recommended for future crystallization studies.

Appendix A

Apparatus and Photochemical Crystal Details

A.1 Physical Parameters of the Photochemical-Heptane System

The heat of crystallization, ΔH_c, and crystal density of the photochemical, ρ_c, were provided by the Eastman Kodak Company. The tabulated values were determined for the photochemical morphology used in this study. The morphology was confirmed using X-ray diffraction techniques.

Physical Parameter	Value	Analysis Technique
ΔH_c	-65 Joules/gram	differential scanning calorimetry
ρ_c	1.183 g/ml	helium displacement

The photochemical solubility in heptane is related to temperature (degrees Celsius) via the following expression.

$$\hat{C}_{\text{sat}}(T) = .185 - 2.11 \times 10^{-2}\ T + 7.46 \times 10^{-4}\ T^2 \tag{A.1}$$

Equation A.1 was determined from concentration data collected during long-term equilibrium experiments in the batch crystallizer apparatus. The correlation is a best fit model to data collected in the $17 - 24°$C temperature range.

A.2 Initial Characterization of Seed Habit

The photochemical seeds were crystallized in the batch crystallizer and separated from the liquor via filtration. After drying, the hardened filter cake was broken, ground with a spatula, and mixed thoroughly to provide a uniform population of seeds. As pictured in the scanning electron micrograph (SEM) of Figure A.1, the seed crystals appear irregular and broken but are approximately parallel-piped in shape. The initial estimates of the average length-to-width and height-to-width ratios of the photochemical seeds were made from Figure A.1 and were later reestimated from optical micrographs (see Figure 3.5).[1] The shape factors were calculated according to the following expressions

$$k_a = 2(l_w d_w + l_w + d_w) \tag{A.2}$$

$$k_v = l_w d_w \tag{A.3}$$

where l_w and d_w are the length-to-width and depth-to-width ratios for the crystals. As measured from 36 crystals in Figure A.1, the average length-to-width ratio of the seed

[1] Length refers to the longest axis or first dimension of a parallel-piped crystal. The second dimension is the width and the third dimension is the depth. Recall that L in the population balance model is crystal width.

crystals was 3.7 with a sample standard deviation of 1.7. The depth-to-width ratio was assumed to be 0.25. These values of l_w and d_w correspond to $k_a = 9.8$ and $k_v = 0.93$.

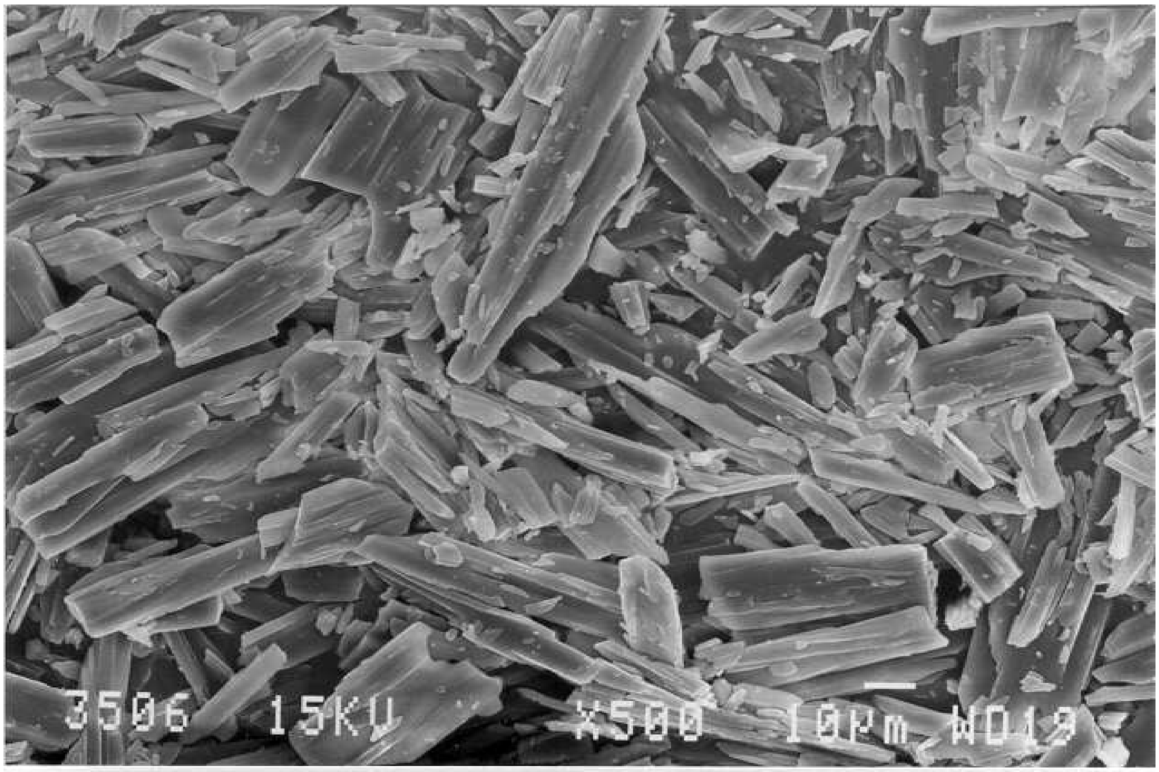

Figure A.1: Scanning electron micrograph of photochemical seed crystals.

A.3 Mettler Settings for the MASSLOG Program

The MASSLOG program collects mass versus time data for the constant filtration tests. The program is designed to work with Mettler BB and PB balances but, under the PB settings, it also collects data from Mettler PM balances. This study used a Mettler PM3000 balance, precise to 0.1 grams.

The MASSLOG program requires the following communication settings on the Mettler PM3000 balance.

Communication Parameter	Value
stability flag	SSTB
baud rate	9600
parity	even

Appendix B

Stochastic Modeling

Despite the fact that stochastic population models are not implemented in this study, a brief examination of the field helps to put the deterministic PBE into perspective. This appendix is meant to motivate the need to understand stochastic models, provide a succinct description of the foundation of such models, and to demonstrate a stochastic formulation for describing the evolution of a single crystal. However, a fundamental description of general stochastic modeling is beyond the scope of this work. The reader seeking a greater understanding of these models is encouraged to consult the texts of van Kampen [150], Gardiner [48], and Grimmett and Stirzaker [60]. Articles specifically related to developments for population modeling are provided by Ramkrishna and Borwanker [117, 118]. The discussion that follows often refers to the modeling of CSD's. However, this is a matter of convenience and such references should be considered surrogates for other types of populations or distributions.

The PBE is a macroscopic approach that describes the time-averaged, observable behavior of a process that is driven by the compilation of a series of unobservable, microscopic events. In crystallization these events may include the addition of a solute molecule into the lattice of a crystal or a collision between two crystals that results in either an agglomerated crystal or a breakage event. Rather than pursue the time-average explicitly, stochastic models are concerned with the random nature of these events. Their use involves the construction of probability functions and expectations to describe how the particular types of events are likely to affect individual crystals. The macroscopic behavior of the system is predicted by averaging over the trajectories of the individual crystals through the state space. For realistic situations stochastic models are rarely solved analytically but are instead used as the basis of Monte Carlo simulation techniques.

In most instances, the macroscopic behavior of a system is adequately described by the PBE. The PBE is adequate if the modeled population is large because microscopic events tend to occur on short time scales relative to those required to obtain a significant change in the macroscopic CSD. In effect, the microscopic events establish an equilibrium that corresponds to the macroscopic state of the system. However, if the population is very small, this equilibrium concept breaks down and modeling the random CSD fluctuations may become important. Ramkrishna and Borwanker [117] show that the relative magnitude of the random fluctuations is $(\sqrt{\mu_0})^{-1}$, where μ_0 is the zeroth moment of the CSD which equals the number of particles in the system.

Statistical Foundations

Stochastic models cope with the time scale problem by modeling the evolution of a particle population not as a time average but as a spatial average over a representative ensemble of individual particles. The location of each particle in the state space evolves from an initial state, or more precisely an initial distribution, due to a series of discrete events (e.g.

molecular collisions). The events constitute realizations, x, of a random variable X. Thus, the location of a particle is a stochastic process, $Y_X(t)$, with a physical realization, $Y_x(t)$. Averages are determined for this process by taking the expectation of $Y_x(t)$

$$\langle Y(t)\rangle = \int Y_x(t)P_X(x)dx \tag{B.1}$$

where $P_X(x)$ is the probability density of X. The probability density for this stochastic process to realize a value y at a given time is

$$P_1(y,t) = \int \delta(y - Y_x)P_X(x)dx. \tag{B.2}$$

The subscript indicates that $P_1(y,t)$ is the density at a single time, t, and distinguishes it from the joint probability density, $P_n(y_1, t_1 : y_2, t_2 : \ldots : y_n, t_n)$, for $Y_x(t)$ to have the value y_1 at t_1, y_2 at t_2, and so on to y_n at t_n. The conditional probability for the process to have the value y at t given that it had the value y' at t' with $t' < t$ is $P_{1|1}(y,t|y',t')$. $P_{1|1}(y,t|y',t')$ is called the transition probability from y', t' to y, t.

The specification of the probability densities of a stochastic process are better pursued via sequential determination of the conditional probabilities. Note that for $t_1 < t_2 < t_3$, we have by the definition of conditional probability [6]

$$P_3(y_3, t_3 : y_2, t_2 : y_1, t_1) = P_{1|2}(y_3, t_3|y_2, t_2 : y_1, t_1)P_2(y_2, t_2 : y_1, t_1) \tag{B.3}$$

Assume that the process has the Markov property.

$$P_{1|n-1}(y_n, t_n|y_1, t_1 : \ldots : y_{n-1}, t_{n-1}) = P_{1|1}(y_n, t_n|y_{n-1}, t_{n-1}) \tag{B.4}$$

The Markov property implies that no historical information prior to time t_{n-1} is required to specify the probability for the state y_{n-1} at t_{n-1} to transition to state y_n at t_n. Thus, a Markov process is fully specified by its initial condition and its transition probability. From these two, all other probabilities may be determined in sequence. With this property $P_{1|2}(y_3, t_3|y_2, t_2 : y_1, t_1) = P_{1|1}(y_3, t_3|y_2, t_2)$ and Equation B.3 may be rewritten

$$\begin{aligned} P_3(y_3, t_3 : y_2, t_2 : y_1, t_1) &= P_{1|1}(y_3, t_3|y_2, t_2) P_2(y_2, t_2 : y_1, t_1) \\ &= P_{1|1}(y_3, t_3|y_2, t_2) P_{1|1}(y_2, t_2|y_1, t_1) P_1(y_1, t_1). \end{aligned} \tag{B.5}$$

Integrating Equation B.5 over all possible values of y_2 and dividing by $P_1(y_1, t_1)$ gives the Chapman-Kolmogorov equation.

$$P_{1|1}(y_3, t_3|y_1, t_1) = \int P_{1|1}(y_3, t_3|y_2, t_2) P_{1|1}(y_2, t_2|y_1, t_1) dy_2. \tag{B.6}$$

The transition probability of any Markov process obeys the Chapman-Kolmogorov equation.

The meaning of the Chapman-Kolmogorov equation is intuitively clear. Imagine the entire time of interest to be $[t_1, t_3]$. Chapman-Kolmogorov relates the probability of the process moving to the final state from the initial state to– 1) the probability that the process may transition out of the initial state into an intermediate state and 2) the probability that the process may transition into the final state from that intermediate state. Thus, if a process cannot transition out of state y_1, the transition probability to $y_3 \neq y_1$ for any t_3 is zero even if y_3 is accessible from an infinite number of intermediate states. The same is true for an inaccessible y_3, regardless of the transient nature of y_1.

By assuming the model to have the Markov property, the probability distribution $P_1(y, t)$ can be determined from the initial distribution, $P_1(y_0, t_0)$, and the transition probability, $P_{1|1}(y, t|y', t')$, and that the Chapman-Kolmogorov equation provides a functional relation that must be satisfied by the transition probability. However, the Chapman-Kolmogorov equation is a general property of all Markov processes and provides no guidance as to the determination of the transition probability for a specific system. To formulate models for specific systems, we resort to the master density equation.

The master density equation is the differential form of the Chapman-Kolmogorov equation for vanishing time difference, Δt. Van Kampen [150] shows that the transition probability for small $\Delta t = t_2 - t_1$ may be written as

$$P_{1|1}(y_2, t_2|y_1, t_1) = (1 - \Delta t \int W(y_2, t_2|y_1, t_1)dy_2)\delta(y_2 - y_1) + \Delta t W(y_2, t_2|y_1, t_1) + O(\Delta t) \tag{B.7}$$

where $W(y_2, t_2|y_1, t_1)$ is the transition probability per unit time from y_1 to y_2 and $O(\Delta t)$ is an error term such that $\lim_{\Delta t \to 0} o(\Delta t)/\Delta t = 0$. Equation B.7 expresses the transition probability as the chance that states $y_1 \neq y_2$ transition into y_2 plus the chance that there is no transition out of y_1 when $y_1 = y_2$. Substituting B.7 into the Chapman-Kolmogorov equation one finds for $\Delta t = t_3 - t_1$

$$\begin{aligned} P_{1|1}(y_3, t_3|y_1, t_1) = & \left(1 - \Delta t \int W(y_2, t_2|y_3, t_3)dy_2\right) P_{1|1}(y_3, t_3|y_1, t_1) \\ & + \Delta t \int W(y_3, t_3|y_2, t_2)P_{1|1}(y_2, t_2|y_1, t_1)dy_2 \end{aligned}$$

Dividing both sides by Δt and taking the limit as $\Delta t \Rightarrow 0$ gives

$$\frac{\partial}{\partial t}P_{1|1}(y_3, t_3|y_1, t_1) = \int \left\{W(y_3, t_3|y_2, t_2)P_{1|1}(y_2, t_2|y_1, t_1) - W(y_2, t_2|y_3, t_3)P_{1|1}(y_3, t_3|y_1, t_1)\right\} dy_2. \tag{B.8}$$

Equation B.8 is the general master density equation for continuous y. If the space of y is discrete, with index n, the master density equation becomes

$$\frac{dp_n(t)}{dt} = \sum_{n \neq n'} \left\{W_{nn'}(t)p_{n'}(t) - W_{n'n}(t)p_n(t)\right\}. \tag{B.9}$$

Clearly, B.9 is a balance equation for the probabilities of the n discrete states. The first term describes all the transitions into a particular state from other states and the second term describes all the transitions out of that state into other states.

The master density equation relates the general stochastic model to the phenomenological details of a specific system.

Example System. Consider a continuous, ideally mixed crystallizer maintained at steady state so that the growth rate is constant and the CSD of the outlet stream is representative of that in the crystallizer. Agglomeration and breakage are neglected. The size coordinate is partitioned into discrete size increments, δ, such that a crystal in the nth state has a size $n\delta$.

During some time interval, $t + \Delta t$, a crystal in the nth state may either grow into another state, wash out of the crystallizer, or remain in state n. Assume Δt to be small such that the following transition probabilities hold

$$\text{Transition by growth} = \alpha\Delta t + o(\Delta t)$$

$$\text{Wash out} = \gamma\Delta t + o(\Delta t)$$

$$\text{Remain in } i = 1 - (\alpha + \gamma)\Delta t + o(\Delta t)$$

and α and γ are the transition intensities of the growth and wash out processes, respectively. Given these transition intensities, the probability that a particular crystal is in state n, for $n \neq 1$, at time $t + \Delta t$ is

$$p_n(t + \Delta t) = [\alpha\Delta t + o(\Delta t)]\, p_{n-1}(t) + [1 - (\alpha + \gamma)\Delta t + o(\Delta t)]\, p_n(t). \tag{B.10}$$

Subtract $p_n(t)$ from both sides, divide by Δt, and take the limit as $\Delta t \Rightarrow 0$.

$$\frac{dp_n(t)}{dt} = \alpha p_{n-1}(t) - (\alpha + \gamma)p_n(t) \tag{B.11}$$

Equation B.11 is the master density equation for p_n where $n \neq 1$. For $n = 1$, there is no growth into the state because there is no smaller size class to grow from, so the master density equation reduces to the second term of B.11.

$$\frac{dp_1(t)}{dt} = -(\alpha + \gamma)p_n(t) \tag{B.12}$$

Subject to the initial condition $p_1(0) = 1$, it may be shown that Equations B.11 and B.12 have the solution

$$p_n(t) = \frac{(\alpha t)^{n-1}}{(n-1)!} \exp\left(-(\alpha + \gamma)t\right) \tag{B.13}$$

which is indicative of the decaying exponential size distribution of a mixed suspension, mixed product removal continuous crystallizer.

The rigorous construction the population balance from a particular master density equation requires the integration of the master density over an ensemble of particles. An example of such an effort for the stochastic simulation of the Boltzmann equation is given by Lymberopoulos and Schieber [91].

Appendix C

Derivation of Relative Supersaturation Expression

The chemical potential difference between the solution and crystalline phases is given as

$$\Delta\mu = \mu_s - \mu_c. \tag{C.1}$$

Through the use of chemical activities each potential may be expressed relative to a standard value.

$$\mu_s = \mu_{\text{std}} + RT\ln(a_s)$$

$$\mu_c = \mu_{\text{std}} + RT\ln(a_c)$$

R is the ideal gas constant and T is the solution temperature. The difference in Equation C.1 is then

$$\Delta\mu = RT\ln\left(\frac{a_s}{a_c}\right). \tag{C.2}$$

At equilibrium, the ratio of activities is equal to one and the solution is saturated with solute. A reasonable approximation is made in the assumption that, even in a nonequilibrium state, the activity of the crystalline phase, a_c, is equivalent to that of the saturated phase, a_{sat}. Thus, the activity ratio may be viewed as ratio of the activities of two liquid solutions. For practical reasons, it is preferable to relate these activities to the solute concentrations in the solutions by resorting to an activity coefficient definition of C.2.

$$\frac{a_s}{a_{\mathrm{sat}}} = \frac{\gamma_s C_s}{\gamma_{\mathrm{sat}} C_{\mathrm{sat}}(T)} \tag{C.3}$$

γ_s and γ_{sat} are the appropriate concentration-dependent activity coefficients. If the concentration difference is relatively small, the ratio of the activity coefficients is taken to be unity and Equation C.2 reduces to

$$\frac{\Delta\mu}{RT} = \ln\left(\frac{C_s}{C_{\mathrm{sat}}}\right) = \ln(1+S) \tag{C.4}$$

where S is the relative supersaturation. S depends on both solution concentration and temperature.

$$S = \frac{C - C_{\mathrm{sat}}(T)}{C_{\mathrm{sat}}(T)} \tag{C.5}$$

Finally, for small S,

$$\ln(1+S) \approx S.$$

Appendix D

Derivation of Mass and Energy Balance Equations

D.1 Continuous phase mass balance

The mass balance for the crystallizing solute in the continuous phase is

$$\frac{d(V_l C)}{dt} = -\sum_k \epsilon_k Q_k C_k + Vr \tag{D.1}$$

where C is the solute concentration in the continuous liquid phase, V_l is the liquid volume, and ϵ_k, Q_k, and C_k are the solids-free volume fraction, volumetric flowrate, and solute concentration in the kth stream, respectively. The sign of Q_k is determined relative to the outward normal of the surface of the control volume, V. The per unit volume reaction term, r, accounts for the generation of continuous phase solute mass via chemical reactions or phase changes. Noting that $V_l = V\epsilon$, Equation D.1 may be recast in terms of the system volume.

$$\frac{d(V\epsilon C)}{dt} = -\sum_k \epsilon_k Q_k C_k + Vr \tag{D.2}$$

Neglecting reactions, Vr is the net rate of solute mass transfer into the continuous phase from the crystalline phase. Clearly, for a closed batch crystallizer

$$Vr = -\frac{dm_c}{dt} \tag{D.3}$$

where m_c is the mass of crystallized solute in the system. dm_c/dt may be determined through an analysis of the population balance equation (PBE) that describes the size distribution of the crystals.

The PBE for the batch cooling crystallizer is

$$\frac{\partial f}{\partial t} = -\frac{\partial (Gf)}{\partial L} + B^\circ \delta(L - L_0) \tag{D.4}$$

The jth moment, μ_j, of the CSD is defined as

$$\mu_j = \int_0^\infty f L^j dL. \tag{D.5}$$

The rate equations for the moments are derived from Equation D.4.

$$\frac{\partial}{\partial t}\int_0^\infty f L^j dL = -\int_0^\infty \frac{\partial (Gf)}{\partial L} L^j dL + \int_0^\infty B^\circ \delta(L - L_0) L^j dL \tag{D.6}$$

Perform integration by parts on the first term on the left hand side of D.6.

$$\begin{aligned} \int_0^\infty \frac{\partial (Gf)}{\partial L} L^j dL &= L^j G f|_0^\infty - j\int_0^\infty G f L^{j-1} dL \\ &= -j\int_0^\infty G f L^{j-1} dL \end{aligned} \tag{D.7}$$

Substitute Equations D.5 and Equation D.7 into Equation D.6 to give an ODE describing the evolution of the jth moment.

$$\frac{d\mu_j}{dt} = j\int_0^\infty G f L^{j-1} dL + B^\circ L_0^j \tag{D.8}$$

Given constant crystal density, ρ_c, and volume shape factor, k_v, the solute mass in the crystalline phase is proportional to the third moment of the CSD.

$$m_c = V\rho_c k_v \mu_3 \tag{D.9}$$

For a constant volume system, the rate of mass transfer into the crystalline phase is

$$\begin{aligned}\frac{dm_c}{dt} &= V\rho_c k_v \frac{d\mu_3}{dt} \\ &= V\left(3\rho_c k_v \int_0^\infty GfL^2 dL + \rho_c k_v B^\circ L_0^3\right)\end{aligned} \tag{D.10}$$

where the parenthetical term on the right hand side is equal to $-r$. Equation D.10 may be substituted into Equation D.2 to give

$$\frac{d(V\epsilon C)}{dt} = -\sum_k \epsilon_k Q_k C_k - V\left(3\rho_c k_v \int_0^\infty GfL^2 dL + \rho_c k_v B^\circ L_0^3\right) \tag{D.11}$$

If $L_0 \Rightarrow 0$ and the stream term is neglected, the final form of the mass balance is a sole function of mass transfer into the solid phase via the crystal growth mechanism.

$$\frac{d(V\epsilon C)}{dt} = -3V\rho_c k_v \int_0^\infty GfL^2 dL \tag{D.12}$$

D.2 Crystallizer energy balance.

The general form of the energy balance is

$$\frac{d}{dt}(U + KE + PE) = -\sum_k \rho_k Q_k \left(\overline{U}_k + \overline{KE}_k + \overline{PE}_k\right) + E_{\text{ext}} + W_t \tag{D.13}$$

where U is internal energy, KE is kinetic energy, PE is potential energy, E_{ext} is energy entering at the system boundary, and W_t is the total work done on the system. The subscript k is indicative of values specific to the kth stream and overbars indicate per unit mass quantities. For systems with appreciable temperature gradients, the kinetic and potential energy terms may be neglected. In addition, W_t may be divided into stream flow and shaft work terms

$$W_t = -\sum_k P_k Q_k + W_s \tag{D.14}$$

where P_k is the pressure of the kth stream at the control volume boundary. Substituting for W_t in Equation D.13 gives

$$\frac{dU}{dt} = -\sum_k \rho_k Q_k \left(\overline{U}_k + \frac{P_k}{\rho_k} \right) + E_{\text{ext}} + W_s. \tag{D.15}$$

In flow systems it is more concise to develop the energy balance in terms of enthalpy, H. Enthalpy is defined as

$$H = U + PV \tag{D.16}$$

and has the per unit mass form

$$\overline{H} = \overline{U} + \frac{P}{\rho}. \tag{D.17}$$

Substituting Equations D.16 and D.17 into D.15 gives

$$\frac{dH(T)}{dt} + \frac{d(PV)}{dt} = -\sum_k \rho_k Q_k \overline{H}_k(T_k) + E_{\text{ext}} + W_s. \tag{D.18}$$

Making the assumption that the system is well-mixed, the H_k's may be referred to the system temperature, T,

$$\begin{aligned} \overline{H}_k(T_k) &= \overline{H}_k(T) + \int_T^{T_k} \frac{\partial H}{\partial T} dT \\ &= \overline{H}_k(T) + \int_T^{T_k} c_{pk} dT \\ &= \overline{H}_k(T) + c_{pk}(T_k - T) \end{aligned} \tag{D.19}$$

where c_{pk} is the heat capacity per unit mass, assumed temperature-independent, of the kth stream. $H_k(T)$ differs from $H(T)$ due solely to composition differences between stream k and the control volume. Substitute Equation D.19 into Equation D.18

$$\frac{dH(T)}{dt} + \frac{d(PV)}{dt} = -\sum_k \rho_k Q_k \left(\overline{H}_k(T) + c_{pk}(T_k - T)\right) + E_{\text{ext}} + W_s. \tag{D.20}$$

Having referred all enthalpies to the system temperature, the parenthetical temperature index is dropped.

Examining the left hand side of D.20, $d(PV)/dt$ may be significant in gas-liquid systems (e.g., evaporative crystallizers) but is negligible for liquid or liquid-solid systems like the cooling crystallizer used in this study. Neglect $d(PV)/dt$ and note that H is a function of system pressure, temperature, and composition.

$$\frac{dH}{dt} = \frac{\partial H}{\partial P}\frac{dP}{dt} + \frac{\partial H}{\partial T}\frac{dT}{dt} + \sum_i \frac{\partial H}{\partial n_i}\frac{dn_i}{dt} \tag{D.21}$$

Of the partial derivative terms, $\partial H/\partial P$ is negligible for liquid-solid systems, $\partial H/\partial T$ is the definition of the total heat capacity of the solution, $\rho V c_p$, and $\partial H/\partial n_i$ is the molar enthalpy of the ith component of the system, $\tilde{H}_i$. Clearly, for the ith component

$$\frac{dn_i}{dt} = -\frac{1}{M_i}\sum_k C_{ik}Q_k + \nu_i V r \tag{D.22}$$

where M_i is the molecular weight of component i, C_{ik} the mass/volume concentration of i in the kth stream, and ν_i the stoichiometric coefficient for the rate of component i production due to reactions or phase changes occurring in the system. With D.22, Equation D.21 becomes

$$\begin{aligned}\frac{dH}{dt} &= \rho V c_p \frac{dT}{dt} - \sum_i \sum_k C_{ik} Q_k \frac{\tilde{H}_i}{M_i} + Vr\left(\sum_i \nu_i \tilde{H}_i\right) \\ &= \rho V c_p \frac{dT}{dt} - \sum_k \sum_i C_{ik} Q_k \overline{H}_i + Vr\Delta H_r\end{aligned} \tag{D.23}$$

where ΔH_r is the "heat of reaction" for r. ΔH_r is negative if r is an exothermic process. Since, the composition changes in the crystallizer are actually phase changes, C_{ik} is actually $\rho_{ik}\epsilon_{ik}$, the product of the density and volume fraction of the ith phase in the kth stream.

Returning to the energy balance, substitute Equation D.23 into Equation D.20, and subtract $Vr\Delta H_r$ from both sides.

$$\rho V c_p \frac{dT}{dt} - \sum_k \sum_i \rho_{ik}\epsilon_{ik} Q_k \overline{H}_i = -\sum_k \sum_i \rho_{ik}\epsilon_{ik} Q_k \overline{H}_{ik} - \sum_k \rho_k Q_k c_{pk}(T_k - T)$$
$$-Vr\Delta H_r + E_{\text{ext}} + W_s. \tag{D.24}$$

Note that the first two terms on the left hand side are equivalent to the first term on the left hand side of Equation D.20. Adding the double sum on the right hand side to both sides gives

$$\rho V c_p \frac{dT}{dt} = -\sum_k \sum_i \rho_{ik}\epsilon_{ik} Q_k \left(\overline{H}_{ik} - \overline{H}_i\right) - \sum_k \rho_k Q_k c_{pk}(T_k - T)$$
$$-Vr\Delta H_r + E_{\text{ext}} + W_s \tag{D.25}$$

The first term in D.25 is due to enthalpy differences that correspond to differences in the relative amounts of solid and liquid phase between the kth stream and the bulk. These enthalpy differences are generally negligible. The second term cannot be discounted for most systems but because the densitometer stream (the only stream in the experimental system) is assumed to return to the batch crystallizer at the same temperature as the bulk, it is identically zero in this case.

To maintain agreement with the mass balance of Section D.1, $-Vr$ is taken to be the rate of solute mass transfer into the crystalline phase and is given by the the right hand side of Equation D.10. Thus, $-\Delta H_r$ is the heat of crystallization, ΔH_c, which is nearly always negative. With these additional simplifications the energy balance reduces to

$$\rho V c_p \frac{dT}{dt} = -V \Delta H_c \left(3\rho_c k_v \int_0^\infty G f L^2 dL + \rho_c k_v B^\circ L_0^3 \right) + E_{\text{ext}} + W_s. \tag{D.26}$$

Allowing $L_0 \Rightarrow 0$, the birth term in D.26 is dropped.

Finally, the batch crystallizer is cooled by a water jacket. The water is assumed to be well-mixed, giving a simple expression for E_{ext}

$$E_{\text{ext}} = UA \left(T_j(t) - T \right) \tag{D.27}$$

where U is the overall heat transfer coefficient, A the surface area available for heat transfer, and $T_j(t)$ the time-dependent temperature of the jacket water. W_s is neglected under the assumption that the energy required to agitate the system is very small relative to jacket and crystallization effects. Thus, the final form of the energy balance for the batch crystallizer is

$$\rho V c_p \frac{dT}{dt} = -V \Delta H_c 3\rho_c k_v \int_0^\infty G f L^2 dL - UA \left(T - T_j(t) \right). \tag{D.28}$$

Appendix E

Derivation of Bayesian Densities and Associated Modal Variances.

E.1 Bayes' Theorem.

Derived from the definition of conditional probability [6],

$$p(\boldsymbol{\psi}, \boldsymbol{Y}) = p(\boldsymbol{\psi}|\boldsymbol{Y})p(\boldsymbol{Y}) = p(\boldsymbol{Y}|\boldsymbol{\psi})p(\boldsymbol{\psi})$$

Bayes Theorem provides a formula for the inference of model parameters, $\boldsymbol{\psi}$, from a given data set, $\boldsymbol{Y}$.

$$p(\boldsymbol{\psi}|\boldsymbol{Y}) = \frac{p(\boldsymbol{Y}|\boldsymbol{\psi})p(\boldsymbol{\psi})}{p(\boldsymbol{Y})} \tag{E.1}$$

$p(\boldsymbol{\psi}|\boldsymbol{Y})$ is the posterior density of the parameters given the data. It summarizes the information about the parameters that is provided both by the data and any knowledge that we may have about the parameters prior to the estimation. Prior information is given by the prior distribution, $p(\boldsymbol{\psi})$. The distribution $p(\boldsymbol{Y})$ is a constant for a given data set and a given model.

$$p(\boldsymbol{Y}) = \int p(\boldsymbol{Y}|\boldsymbol{\psi})p(\boldsymbol{\psi})d\boldsymbol{\psi}$$

Fisher coined the term "likelihood function" to describe $p(\boldsymbol{Y}|\boldsymbol{\psi})$ when $\boldsymbol{Y}$ is given. The likelihood is written as $\ell(\boldsymbol{\psi}|\boldsymbol{Y})$, giving the form of Equation E.1 most often associated with parameter estimation.

$$p(\boldsymbol{\psi}|\boldsymbol{Y}) \propto \ell(\boldsymbol{\psi}|\boldsymbol{Y})p(\boldsymbol{\psi}) \tag{E.2}$$

$$\text{Posterior} \propto \text{Data} \times \text{Prior}$$

Bayesian estimation has the pleasant interpretation that the parameters are inferred using all previous knowledge plus the new information contained in $\boldsymbol{Y}$. Note that the solution to the Bayesian estimation problem is not a set of point estimates but a full distribution. The optimal parameter estimates are the values that select the maximum, or mode, of this distribution. Thus, these parameters are often termed the "modal" estimates. The uncertainty in the parameters is a function of the spread of the distribution.

E.2 The Likelihood Function.

$p(\boldsymbol{Y}|\boldsymbol{\psi})$ is the joint probability density of the collected data that make up $\boldsymbol{Y}$, given the parameters $\boldsymbol{\psi}$. The model predicts $\boldsymbol{Y}$, subject to a subset of $\boldsymbol{\psi}$.

$$y_i = \tilde{y}_i(\boldsymbol{\theta}) + e_i \tag{E.3}$$

where y_i is the ith data point, $\boldsymbol{\theta}$ the model parameter subset of $\boldsymbol{\psi}$, $\tilde{y}_i(\boldsymbol{\theta})$ the model prediction, and e_i the prediction error. If the ith error is normally distributed with a given measurement variance, σ_i, the probability distribution of y_i is

$$p(y_i|\boldsymbol{\theta}, \sigma_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{(y_i - \tilde{y}_i(\boldsymbol{\theta}))^2}{2\sigma_i^2}\right). \tag{E.4}$$

If the errors are assumed to be independent, we have

$$p(\boldsymbol{Y}|\boldsymbol{\theta}, \boldsymbol{\sigma}) \approx \prod_{i=1}^{n} p(y_i|\boldsymbol{\theta}, \sigma_i) \tag{E.5}$$

where n is the number of individual data points in $\boldsymbol{Y}$ and $\boldsymbol{\sigma}$ is the n-dimensional vector of measurement standard deviations. The independence assumption sets all error correlations to zero, including time correlations and correlations between different types of sensors. Combining Equations E.4 and E.5 we have

$$p(\boldsymbol{Y}|\boldsymbol{\theta}, \boldsymbol{\sigma}) = (2\pi)^{-n/2} \prod_{i=1}^{n} \left[\sigma_i^{-1} \exp\left(-\frac{e_i^2(\boldsymbol{\theta})}{2\sigma_i^2}\right)\right] \tag{E.6}$$

where $e_i(\boldsymbol{\theta})$ is substituted for $(y_i - \tilde{y}_i(\boldsymbol{\theta}))$. Adding the additional simplification that all data collected from a particular type of sensor have a common variance, E.6 becomes

$$p(\boldsymbol{Y}|\boldsymbol{\theta}, \boldsymbol{\sigma}) = \left(\sqrt{2\pi}\right)^{-\left(\sum_{k=1}^{m} n_k\right)} \prod_{k=1}^{m} \sigma_k^{-n_k} \exp\left(-\frac{1}{2}\sum_{k=1}^{m} \sigma_k^{-2} \sum_{i=1}^{n_k} e_{ik}^2(\boldsymbol{\theta})\right) \tag{E.7}$$

where there are m sensor types, n_k is the number of data points available from the kth sensor, and $e_{ik}(\boldsymbol{\theta})$ is the prediction error for the ith measurement from the kth sensor. Clearly, $\boldsymbol{\sigma}$ is now m-dimensional. For given $\boldsymbol{Y}$ as opposed to given $\boldsymbol{\theta}$ and $\boldsymbol{\sigma}$, Equation E.7 is the likelihood function, $\ell(\boldsymbol{\theta}, \boldsymbol{\sigma}|\boldsymbol{Y})$.

Further derivations are simplified by utilizing the natural logarithm of $\ell(\boldsymbol{\theta}, \boldsymbol{\sigma}|\boldsymbol{Y})$. Because the logarithm is a monotonic function of its argument, the simplification is justified for treatments of the modal region of the likelihood function. Taking the logarithm of $\ell(\boldsymbol{\theta}, \boldsymbol{\sigma}|\boldsymbol{Y})$ and designating this quantity by $\mathcal{L}$ gives

$$\mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\sigma}|\boldsymbol{Y}) = -\frac{1}{2}\left(\sum_{k=1}^{m} n_k\right) \ln 2\pi - \sum_{k=1}^{m} n_k \ln \sigma_k - \frac{1}{2}\sum_{k=1}^{m} \sigma_k^{-2} \sum_{i=1}^{n_k} e_{ik}^2(\boldsymbol{\theta}). \tag{E.8}$$

Examining equation E.8 it is instructive to note that if the σ_k's are available *a priori* the $\boldsymbol{\theta}$ estimates that maximize this form of the likelihood function also minimize the summations in the last term.

$$\Phi_{\mathrm{wls}} = \sum_{k=1}^{m} \sigma_k^{-2} \sum_{\mu=1}^{n_k} e_{ik}^2(\boldsymbol{\theta}) \tag{E.9}$$

Φ_{wls} is the objective function for weighted least squares parameter estimation. Weighted least squares is the maximization of the likelihood function given known measurement variances.

If the σ_k's are not available *a priori* they may be estimated explicitly or expressions may be derived that allow them to be estimated from the modal values of the model parameters, $\hat{\boldsymbol{\theta}}$. Take the partial derivative of Equation E.8 with respect to any σ_k and set it equal to zero.

$$\frac{\partial \mathcal{L}}{\partial \sigma_k} = -\frac{n_k}{\sigma_k} + \frac{\sum_{i=1}^{n_k} e_{ik}^2(\boldsymbol{\theta})}{\sigma_k^3} = 0$$

Solving for the optimal value of σ_k^2 gives

$$\hat{\sigma}_k^2(\boldsymbol{\theta}) = \frac{\sum_{i=1}^{n_k} e_{ik}^2(\boldsymbol{\theta})}{n_k}. \tag{E.10}$$

Note that $\hat{\sigma}_k^2$ is strictly positive and attributes all estimation error to variance in the kth sensor measurement. Evaluating the second partial derivative for $\sigma_k = \hat{\sigma}_k$ confirms that $\hat{\sigma}_k$ maximizes $\mathcal{L}$.

$$\begin{aligned}\left(\frac{\partial^2 \mathcal{L}}{\partial {\sigma_k}^2}\right)_{\hat{\sigma}_k} &= \frac{n_k}{\hat{\sigma}_k^2} - \frac{3\sum_{i=1}^{n_k} e_{in}^2(\boldsymbol{\theta})}{\hat{\sigma}_k^4} \\ &= \frac{n_k}{\hat{\sigma}_k^2} - \frac{3n_k}{\hat{\sigma}_k^2} < 0 \end{aligned} \tag{E.11}$$

Substituting $\hat{\sigma}_k$ into Equation E.8, the first and last terms are constant and the maximization of the likelihood function reduces to the minimization of

$$\Phi_{\text{mle}} = \sum_{k=1}^{m} n_k \ln \left(\sum_{i=1}^{n_k} e_{ik}^2(\boldsymbol{\theta}) \right). \tag{E.12}$$

Similar to weighted least squares, the inference of $\boldsymbol{\theta}$ via the minimization of Equation E.12 is a special case of maximum-likelihood estimation. Bard [4] provides an extensive treatment of maximum-likelihood estimation techniques.

E.3 The Noninformative Prior.

Bayes theorem incorporates *a priori* information through the use of a prior distribution on the parameters. However, this concept can be troublesome. For instance, if the prior information is reliable, there is generally no need to continue the investigation, while if the prior information is not reliable, it seems inappropriate to allow it to bias the estimates. If one has no initial knowledge, how can any prior make sense? One solution to this quandary is the use of Jeffrey's noninformative prior [78].

The idea of the Jeffreys prior is intuitive. Recall Equation E.2 and note that $p(\boldsymbol{Y}|\boldsymbol{\psi})$ is multiplied by the prior distribution. Jeffreys sought a prior that would make this product a *data-translated* distribution. Data-translated implies that different $\boldsymbol{Y}$'s position the distribution differently with respect to the parameter space but that the shape of the distribution is unaffected by such relocation (i.e. the shape is not a function of $\boldsymbol{\psi}$). If the shape is not independent of $\boldsymbol{\psi}$, then a particular parameterization, $\boldsymbol{\phi}(\boldsymbol{\psi})$, is derived for which the independence is preserved. From this reasoning, Jeffrey's derived the following general equation for his prior,

$$p(\psi) \propto \sqrt{|\boldsymbol{I}_n|} \tag{E.13}$$

where $\boldsymbol{I}_n$ is the Fisher information matrix defined as

$$\boldsymbol{I}_n = -E\left(\frac{\partial^2 \ln p(\boldsymbol{Y}|\boldsymbol{\psi})}{\partial\psi_i\partial\psi_j}\right). \tag{E.14}$$

The information matrix is the negative of the expected modal curvature of $p(\boldsymbol{Y}|\boldsymbol{\psi})$. Therefore, the Jeffreys prior recasts the independence of shape as the independence of modal curvature.

Given the development of the likelihood function in the previous section, the parameter vector $\boldsymbol{\psi}$ consists of the model parameters, $\boldsymbol{\theta}$, and the m-dimensional measurement standard deviation vector, $\boldsymbol{\sigma}$. It is reasonable to assume that $\boldsymbol{\theta}$ and $\boldsymbol{\sigma}$ are independently distributed such that

$$p(\boldsymbol{\psi}) \approx p(\boldsymbol{\theta})p(\boldsymbol{\sigma}). \tag{E.15}$$

$\boldsymbol{\theta}$ are called *location parameters* because their optimal values are determined from the location of the mode of $p(\boldsymbol{\theta}, \boldsymbol{\sigma}|\boldsymbol{Y})$. Similarly, $\boldsymbol{\sigma}$ are called *scale parameters* because their modal values affect the spread of the distribution. For location parameters, no reparameterization is necessary as the Jeffreys prior generally suggests a locally uniform distribution.[1]

$$p(\boldsymbol{\theta}) = \text{constant} \tag{E.16}$$

Scale parameters, however, do require reparameterization.

Determining Jeffreys' prior for $\boldsymbol{\sigma}$, requires the identification of the elements of the corresponding information matrix. Equation E.11 gives the expected curvature of $\mathcal{L}$ in terms of the modal value of the kth element of $\boldsymbol{\sigma}$. Substituting this expression into Equation E.14, the expected value of k, k element of $\boldsymbol{I}_n$ is

$$\boldsymbol{I}_n(k,k) = \frac{2n_k}{\sigma_k^2}.$$

[1]The distribution is defined as local to allow normalization to unit probability.

Noting that the off-diagonal terms of the $m \times m$ information matrix are zero due to independence assumptions, the determinant of $\boldsymbol{I}_n$ is

$$|\boldsymbol{I}_n| = \prod_{k=1}^{m} \frac{2n_k}{\sigma_k^2}. \tag{E.17}$$

The corresponding prior distribution for $\boldsymbol{\sigma}$ is given by substituting E.17 into Equation E.13.

$$p(\boldsymbol{\sigma}) \propto \prod_{k=1}^{m} \sigma_k^{-1}. \tag{E.18}$$

Given Equation E.16, $p(\boldsymbol{\psi}) \propto p(\boldsymbol{\sigma})$.

Box and Tiao [16] and Stewart et al. [141] provide extensive discussions regarding the origins and use of Jeffreys prior.

E.4 The Posterior Distribution, Modal Measurement Variance, and Objective Function.

Given that $p(\boldsymbol{\theta})$ is locally uniform, Equations E.7 and E.18 may be substituted into Equation E.2 to give the posterior distribution appropriate for the assumptions of Section E.2.

$$p(\boldsymbol{\theta}, \boldsymbol{\sigma}|\boldsymbol{Y}) \propto \left(\sqrt{2\pi}\right)^{-\left(\sum_{k=1}^{m} n_k\right)} \prod_{k=1}^{m} \sigma_k^{-(n_k+1)} \exp\left(-\frac{1}{2}\sum_{k=1}^{m} \sigma_k^{-2} \sum_{i=1}^{n_k} e_{ik}^2(\boldsymbol{\theta})\right) \tag{E.19}$$

By taking the natural logarithm of $p(\boldsymbol{\theta}, \boldsymbol{\sigma}|\boldsymbol{Y})$ and proceeding with a development similar to that performed for the likelihood function in Section E.2, it may be shown that the modal parameters of $p(\boldsymbol{\theta}, \boldsymbol{\sigma}|\boldsymbol{Y})$ are equivalent to those minimizing the following objective function.

$$\Phi_{\text{be}} = \sum_{k=1}^{m} (n_k + 1) \ln \left(\sum_{i=1}^{n_k} e_{ik}^2(\boldsymbol{\theta})\right). \tag{E.20}$$

Equation E.20 corresponds to the following modal estimate of the m measurement variances.

$$\hat{\sigma}_k^2 = \frac{\sum_{i=1}^{n_k} e_{ik}^2(\hat{\boldsymbol{\theta}})}{n_k + 1} \tag{E.21}$$

The posterior distribution corresponding to Φ_{be} is written $p(\boldsymbol{\theta}, \hat{\boldsymbol{\sigma}}|\boldsymbol{Y})$ to denote the use of the modal value of $\boldsymbol{\sigma}$.

Appendix F

Comments on Transmittance Data and Dissolution.

F.1 Estimation of Parameters without Transmittance Data

The impact of the transmittance data on the estimates may be assessed by removing this data from the estimation data set and reestimating the parameters using only concentration and microscope data. The optimal estimates determined for the concentration and microscope data are given in Table F.1. In general, the changes in the optimal values of the parameters are insignificant. However, the optimal estimates for $\ln(k_g)$, $L_{\min}$ and $l_w(\infty)$ do fall outside of the linear 95% confidence intervals determined for their full data set estimates (given in Table 4.11. The parameter most effected by the removal of the transmittance data was $l_w(\infty)$.

The exclusion of transmittance data had a noticeable effect on the interval of $l_w(\infty)$. The width of the $l_w(\infty)$ confidence interval increased 58%. The parameters least effected by the removal of the transmittance data were the growth rate parameters and the third

	$\ln(k_g)$	g	$\ln(k_b)$	b	j	$L_{\min}$	$l_w(\infty)$	k_f
Estimate	-1.43	1.75	27.16	2.88	2.21	26.46	12.40	0.40
Interval	±0.04	±0.02	±0.59	±0.19	±0.08	±0.73	±1.28	±0.03

Table F.1: Parameter estimates and 95% confidence intervals for the optimal disturbance model excluding all transmittance data.

moment order of nucleation, j.

The fits of the model described in Table F.1 to the transmittance data are given in Figures F.1 and F.2. The degradation of the fits relative to those of the optimal disturbance model (see Figures 4.37 through 4.41) is minimal but is most apparent in run 1c and the A-Optimal experiment.

The importance of the transmittance measurement for the estimation of precise parameters is illustrated by the increased uncertainty in the $l_w(\infty)$ estimate. However, because the estimates in Table F.1 provide satisfactory fits to the transmittance data and are reasonably precise it may appear that transmittance measurements were not necessary to identify the kinetics of this system. However, this interpretation is not correct. Recall that in the initial stage of the parameter estimation study, large transmittance residuals drew attention to errors in the model and data. These errors were not detectable in the small concentration residuals but were later confirmed experimentally (see Sections 4.5.3 and 4.9). The model and data errors would not have been identified or even investigated had transmittance data not been collected. Thus, despite the relatively small impact of the transmittance data on the optimal estimates, transmittance data played a crucial role in the model identification stage of this study.

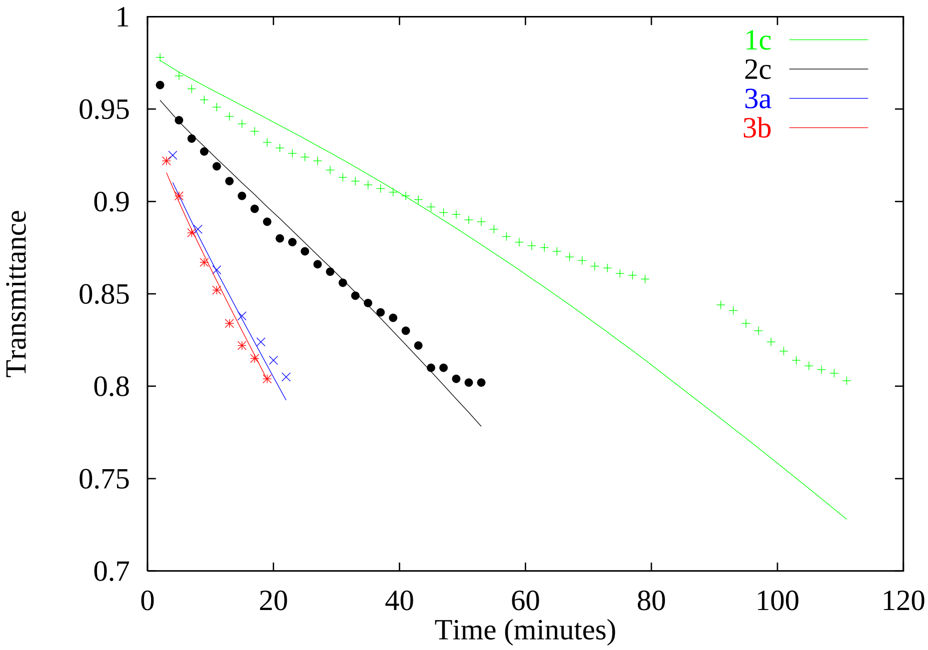

Figure F.1: Transmittance profiles and model fits for runs 1c, 2c, 3a, and 3b using the parameters in Table F.1.

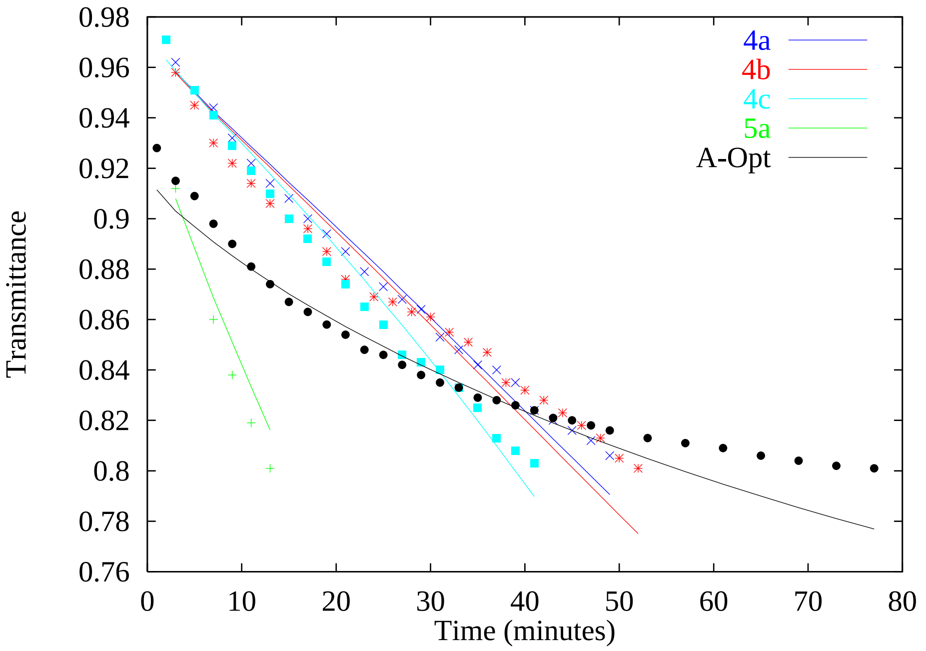

Figure F.2: Transmittance profiles and model fits for runs 4a, 4b, 4c, 5a, and A-Opt using the parameters in Table F.1.

F.2 Dissolution Experiments.

An initial goal of this study was to determine the dissolution kinetics of the photochemical-heptane system. As noted in Section 3.3.3, adding dissolution phases to standard crystallizations may provide crystal shape or size distribution improvements that will facilitate downstream solid-liquid separation processes. This section describes a simple experiment that was performed to demonstrate why the dissolution kinetics for the photochemical-heptane system could not be determined in this study.

In order to produce an initial population of crystals for the dissolution experiment, crystallization experiment similar to the run 4 replicates (recall Table 4.1) was conducted for a period of 250 minutes. This allowed the dissolution phase model to be initialized with the state values from the crystallization phase and assured that the dissolution experiment would be conducted on a crystal population with the appropriate habit and morphology. After 250 minutes, the temperature in the system was put through a series of ramped temperature elevations. The concentration data collected during this ramping phase are given with the corresponding temperature profile in Figure F.3.

Figure F.3 shows that, unlike the slow process of crystal growth, the dissolution of crystals in the photochemical-heptane system is very fast. The concentration data follow the 305 and 340 minute temperature ramps with very little delay and reach equilibrium within 2-3 minutes after the temperature becomes constant again. The speed of the process negatively impacted the estimation of the dissolution parameters by limiting both the amount of dynamic data that could be collected and the driving force level that could be achieved during an experiment. The dissolution process was sufficiently fast that it was

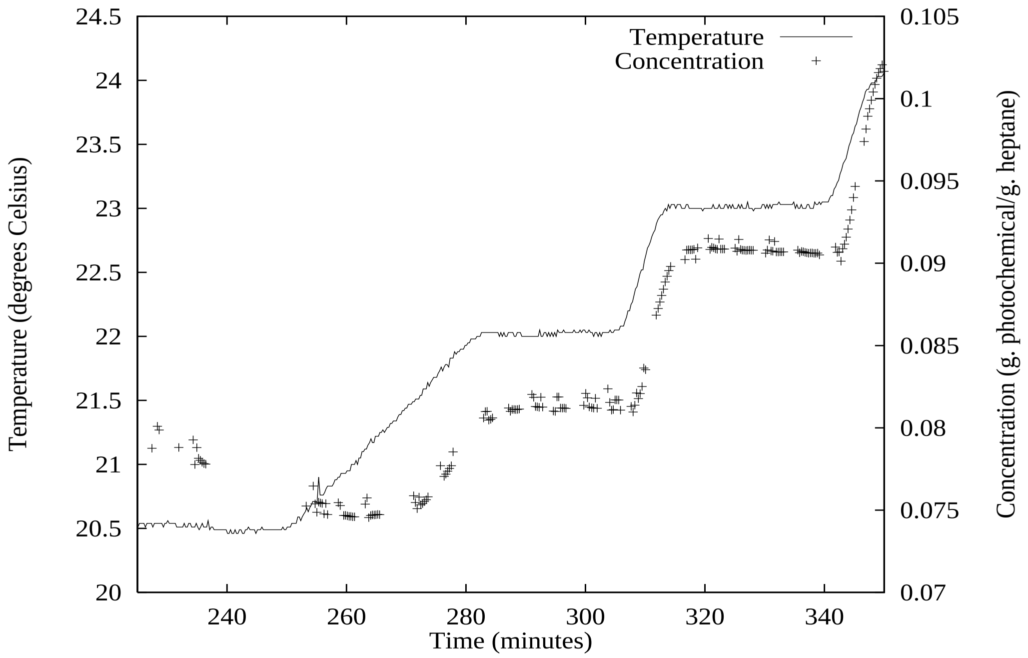

Figure F.3: Concentration and temperature profile for the dissolution phase.

not possible for the apparatus to achieve supersaturations levels lower than -0.05. This contrasted sharply with the crystallization experiments that had driving forces as high as 1.0.

The low driving forces also exposed errors in the solubility curve. This is illustrated by attempting to fit the dissolution data with the dissolution rate model of Equation 3.22a. Such a fit is shown in Figure F.4. Note the mismatch between the model and the data during the equilibrium periods. These prediction errors resulted from errors in the solubility curve. These same errors existed in the crystallization experiments but their impact was less of a concern because the errors were generally small relative to the driving forces used. With the small driving forces of dissolution, the solubility curve errors become relatively substantial. Clearly, the dissolution parameters used to fit the data in Figure F.4 are meaningless given the solubility errors.

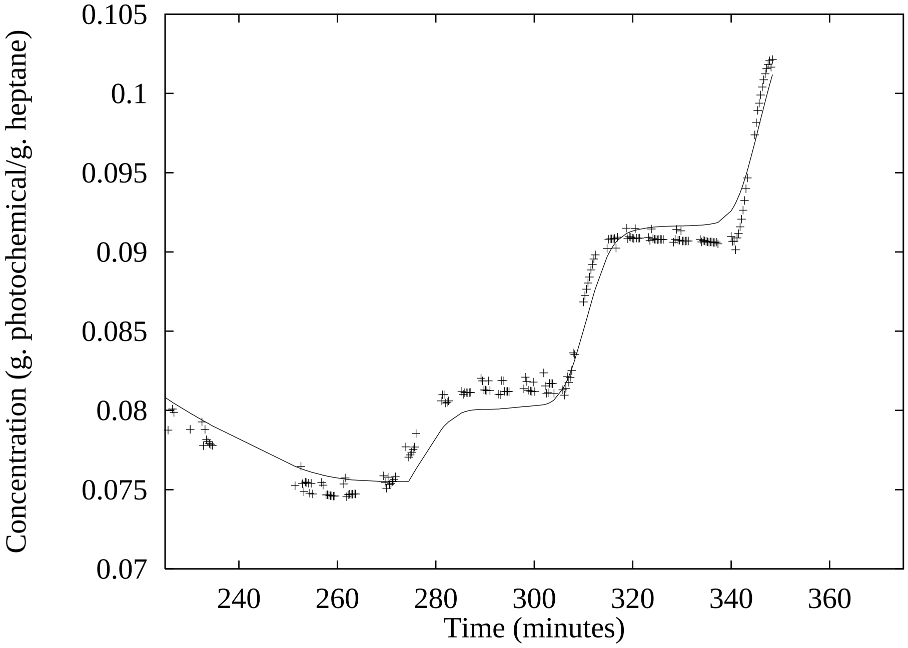

Figure F.4: Model fit to crystallization/dissolution data.

The failure to estimate dissolution kinetics in the photochemical-heptane system, while unfortunate in an academic sense, is not particularly troubling from a practical standpoint. Put simply, the process is too fast for its dynamics to be exploited in an optimal control strategy. If a dissolution phase is desired, the temperature may be set to correspond to the equilibrium concentration desired by the investigator and the system may be assumed to equilibrate almost immediately.

Bibliography

[1] C. F. Abegg, J. D. Stevens, and M. A. Larson. Crystal size distributions in continuous crystallizers when growth rate is size dependent. *AIChE Journal*, 14(1):118–122, January 1968.

[2] M. B. Ajinkya and W. H. Ray. On the optimal operation of crystallization processes. *Chemical Engineering Communications*, 1:181–186, 1974.

[3] A. C. Atkinson and A. N. Donev. *Optimum Experimental Designs*. Oxford University Press, New York, 1992.

[4] Yonathan Bard. *Nonlinear Parameter Estimation*. Academic Press, New York, 1974.

[5] L. G. Bauer, M. A. Larson, and V. J. Dallons. Contact nucleation of $mgso_4 \cdot 7h_2o$ in a continuous msmpr crystallizer. *Chemical Engineering Science*, 29:1253–1261, 1974.

[6] Thomas Bayes. An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London*, 53:370–418, 1763. Reprinted in *Biometrika*, 35:293–315, 1958.

[7] D. Behnken, J. Horowitz, and S. Katz. Particle growth processes. *Industrial and Engineering Chemistry Fundamentals*, 2(3):212–216, August 1963.

[8] H. L. Bhat, J. N. Sherwood, and T. Shripathi. The influence of stress, strain and fracture of crystals on the crystal growth process. *Chemical Engineering Science*, 42(4):609–618, 1987.

[9] H. L. Bhat, J. N. Sherwood, and T. Shripathi. Strain variation in the { 100} growth sectors of potash alum single crystals and its relationship to growth rate dispersion. *Journal of Crystal Growth*, 102:245–248, 1990.

[10] M. Bohlin and Å. C. Rasmuson. Application of controlled cooling and seeding in batch crystallization. *The Canadian Journal of Chemical Engineering*, 70:120–126, 1992.

[11] Gregory D. Botsaris. Secondary nucleation—a review. In J. W. Mullin, editor, *Industrial Crystallization*, New York, 1976. Plenum Press.

[12] G. E. B. Box and H. L. Lucas. Design of experiments in non-linear situations. *Biometrika*, 46:77–90, 1959.

[13] G. E. P. Box and W. J. Hill. Discrimination among mechanistic models. *Technometrics*, 9(1):57–71, 1967.

[14] G. E. P. Box and W. J. Hunter. Sequential design of experiments for nonlinear models. In *IBM Scientific Computing Symposium in Statistics*, pages 113–137. IBM, 1965.

[15] George E. P. Box, William G. Hunter, and J. Stuart Hunter. *Statistics for Experimenters.* John Wiley & Sons, New York, 1978.

[16] George E. P. Box and George C. Tiao. *Bayesian Inference in Statistical Analysis.* Addison–Wesley Publishing Company, Reading, Massachusetts, 1st edition, 1973.

[17] Arthur Boxman. *Particle Size Measurement for Control of Industrial Crystallizers.* PhD thesis, Delft University of Technology, 1992.

[18] D. J. Brown and P. G. Felton. Direct measurement of concentration and size for particles of different shapes using laser light diffraction. *Chemical Engineering Research and Design*, 63:125–132, 1985.

[19] J. Budz, P. H. Karpinski, and Z. Nuruc. Influence of hydrodynamics on crystal growth and dissolution in a fluidized bed. *AIChE Journal*, 30(5):710–717, September 1984.

[20] J. Budz, P. H. Karpinski, and Z. Nuruc. Effect of temperature on crystallization and dissolution processes in a fluidized bed. *AIChE Journal*, 31(2):259–268, February 1985.

[21] W. K. Burton, N. Cabrera, and F. C. Frank. The growth of crystals and the equilibrium structure of their surfaces. *Philosophical Transactions of the Royal Society of London*, 243:299–358, June 1951.

[22] T. F. Canning and A. D. Randolph. Some aspects of crystallization theory: Systems that violate McCabe's delta L law. *AIChE Journal*, 13(1):5–10, 1967.

[23] N. W. Cayey and J. Estrin. Secondary nucleation in agitate magnesium sulfate solutions. *Industrial and Engineering Chemistry Fundamentals*, 6(1):13–20, 1967.

[24] R. Chakraborty and K. A. Berglund. Steady state fluorescence spectroscopy of pyranine as a trace extrinsic probe to structure in aqueous sugar solutions. *Journal of Crystal Growth*, 125:81–96, 1992.

[25] Chuei-Tin Chang and Mary A. Farrell Epstein. Simulation studies of a feedback control strategy for batch crystallizers. *AIChE Symposium Series*, 83(253):110–119, 1987.

[26] R. Chang and M. Wang. Shifted Legendre function approximation of differential equations: Application to crystallization processes. *Computers & Chemical Engineering*, 8(2):117–125, 1984.

[27] Ming Chen and M. A. Larson. Crystallization kinetics of calcium nitrate tetrahydrate from MSMPR crystallizer. *Chemical Engineering Science*, 40(7):1287–1294, 1985.

[28] A. Chianese, A. Condò, and B. Mazzarotta. The growth and dissolution of sodium perborate crystals in a fluidized bed crystallizer. *Journal of Crystal Growth*, 97:375–386, 1989.

[29] N. A. Clontz and W. L. McCabe. Contact nucleation of magnesium sulfate heptahydrate. *AIChE Symposium Series*, 67(110):6–17, 1971.

[30] R. Curl. Dispersed phase mixing: I. Theory and effect in simple reactors. *AIChE Journal*, 9:175, March 1963.

[31] J. E. Cuthrell and L. T. Biegler. On the optimization of differential-algebraic systems. *AIChE Journal*, 33(8):1257–1270, 1987.

[32] S. R. Dash and S. Rohani. Iterative parameter estimation for extraction of crystallization kinetics of potassium chloride from batch experiments. *The Canadian Journal of Chemical Engineering*, 71:539–548, 1993.

[33] P. J. Daudey, G. M. van Rosmalen, and E. J. De Jong. Secondary nucleation kinetics of ammonium sulphate in CMSMPR crystallizer. *Journal of Crystal Growth*, 99:1076–1081, 1990.

[34] R. David, J. Villermaux, P. Marchal, and J. P. Klein. Crystallization and precipitation engineering—4. kinetic model of adipic acid crystallization. *Chemical Engineering Science*, 46(4):1129–1136, 1991.

[35] Dilum D. Dunuwila and Kris A. Berglund. Atr ftir spectroscopy for *in situ* measurement of supersaturation. Submitted for publication in *Journal of Crystal Growth*, May 1996.

[36] Dilum D. Dunuwila, Leslie B. Carroll II, and Kris A. Berglund. An investigation of the applicability of attenuated total reflection infrared spectroscopy for measurement of solubility and supersaturation of aqueous citric acid solutions. *Journal of Crystal Growth*, 137:561–568, 1994.

[37] R. A. Eek, S. Dijkstra, and G. M. van Rosmalen. Dynamic modeling of suspension crystallizers using experimental data. *AIChE Journal*, 41(3):571–584, March 1995.

[38] D. Espie and S. Macchietto. The optimal design of dynamic experiments. *AIChE Journal*, 35(2):223–229, February 1989.

[39] R. J. Farrell and Y. Tsai. Modeling, simulation, and kinetic parameter estimation in batch crystallization processes. *AIChE Journal*, 40(4):586–593, April 1994.

[40] R. J. Farrell and Y. Tsai. Nonlinear controller for batch crystallization: Development and experimental demonstration. *AIChE Journal*, 41:2318–2321, 1995.

[41] V. V. Fedorov. *Theory of Optimal Experiments.* Academic Press, New York and London, 1972.

[42] G. Févotte and J. P. Klein. Application of on-line calorimetry to the advanced control of batch crystallizers. *Chemical Engineering Science*, 49(9):1323–1336, 1994.

[43] Thomas J. Fitzgerald and Tai-Cheng Yang. Size distribution for crystallization with continuous growth and breakage. *Industrial and Engineering Chemistry Fundamentals*, 11(4):588–590, 1972.

[44] Mary Flohr, 1997. Personal Communication.

[45] Andrew P. Fordyce and James B. Rawlings. Modelling antibiotic production using a structured/segregated approach. AIChE National Meeting, Chicago, IL, November 1990.

[46] H. Garabedian and R. F. Strickland-Constable. Collision breeding of crystal nuclei: sodium chlorate. i. *Journal of Crystal Growth*, 13/14:506–509, 1972.

[47] Carlos E. García and A. M. Morshedi. Quadratic programming solution of dynamic matrix control (QDMC). *Chemical Engineering Communications*, 46:73–87, 1986.

[48] Crispin W. Gardiner. *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences.* Springer-Verlag, Berlin, Germany, 2nd edition, 1990.

[49] J. Garside, R. Janssen-van Rosmalen, and P. Bennema. Verification of crystal growth rate equations. *Journal of Crystal Growth*, 29:353–366, 1975.

[50] J. Garside and J. W. Mullin. Continuous measurement of solution concentration in a crystalliser. *Chemistry and Industry*, pages 2007–2008, November 1966.

[51] J. Garside, I. T. Rusli, and M. A. Larson. Origin and size distribution of secondary nuclei. *AIChE Journal*, 25(1):57–64, 1979.

[52] John Garside. Advances in the characterisation of crystal growth. *AIChE Symposium Series*, 80(240):23–38, 1984.

[53] John Garside, V. Roger Phillips, and Mukund B. Shah. On size-dependent crystal growth. *Industrial and Engineering Chemistry Fundamentals*, 15(3):230–233, 1976.

[54] John Garside and Mukund B. Shah. Crystallization kinetics from MSMPR crystallizers. *Industrial and Engineering Chemistry Process Design and Development*, 19:509–514, 1980.

[55] Philip E. Gill, Walter Murray, Michael A. Saunders, and Margaret H. Wright. User's guide for SOL/NPSOL (Version 4.0): A Fortran package for nonlinear programming, technical report SOL 86-2. Technical report, Systems Optimization Laboratory, Department of Operations Research, Stanford University, 1986.

[56] M. W. Girolami and R. W. Rousseau. Size-dependent crystal growth—a manifestation of growth rate dispersion in the potassium alum-water system. *AIChE Journal*, 31(11):1821–1828, 1985.

[57] J. R. Gooch and M. J. Hounslow. Monte Carlo simulation of size-enlargement mechanisms in crystallization. *AIChE Journal*, 42(7):1864–1874, 1996.

[58] Daniel A. Green, 1988. November/December Du Pont Magazine.

[59] Hugh Griffiths. Mechanical crystallisation. *Journal of The Society of Chemical Industry Transactions*, 44:7T–18T, 1925.

[60] G. R. Grimmett and D. R. Stirzaker. *Probability and Random Processes*. Oxford University Press, New York, 2nd edition, 1992.

[61] R. W. Hartel, K. A. Berglund, S. M. Gwynn, P. M. Schierholz, and V. G. Murphy. Crystallization kinetics for the sucrose–water system. *AIChE Symposium Series*, 76(193):65–72, 1980.

[62] S. K. Heffels and E. J. de Jong. Improved operation and control of batch crystallizers. *AIChE Symposium Series*, 87(284):170–181, 1991.

[63] James E. Helt and Maurice A. Larson. Effect of temperature on the crystallization of potassium nitrate by direct measurement of supersaturation. *AIChE Journal*, 23(6):822–830, 1977.

[64] Ladislav Hložny, Akira Sato, and Noriaki Kubota. On-line measurement of supersaturation during batch cooling crystallization of ammonium alum. *Journal of Chemical Engineering of Japan*, 25(5):604–606, 1992.

[65] L. H. Hosten. A sequential experimental design procedure for precise parameter estimation based upon the shape of the joint confidence region. *Chemical Engineering Science*, 29:2247–2252, 1974.

[66] L. H. Hosten and G. Emig. Sequential experimental design procedures for precise parameter estimation in ordinary differential equations. *Chemical Engineering Science*, 30:1357–1364, 1975.

[67] M. J. Hounslow, R. L. Ryall, and V. R. Marshall. A discretized population balance for nucleation, growth, and agglomeration. *AIChE Journal*, 34(11):1821–1832, November 1988.

[68] J. Hsu and L. Fan. Transient analysis of crystallization: Effect of the size dependent growth rate. *Chemical Engineering Communications*, 56:19–40, 1987.

[69] J. P. Hsu, L. T. Fan, and S. T. Chou. Transient analysis of crystallization: Effect of the initial size distribution. *Chemical Engineering Communications*, 69:95–114, 1988.

[70] H. M. Hulbert and Tetsuo Akiyama. Liouville equations for agglomeration an dispersion processes. *Industrial and Engineering Chemistry Fundamentals*, 8(2):319–324, 1969.

[71] H. M. Hulbert and S. Katz. Some problems in particle technology: A statistical mechanical formulation. *Chemical Engineering Science*, 19:555–574, 1964.

[72] W. G. Hunter and A. M. Reiner. Designs for discriminating between two rival models. *Technometrics*, 7(3):307–323, 1965.

[73] D. Ilievski and M. J. Hounslow. Agglomeration during precipitation—II. mechanism deduction from tracer data. *AIChE Journal*, 41(3):525–535, March 1995.

[74] J. Jager, S. de Wolf, W. Klapwijk, and E. J. de Jong. A new design for on-line product-slurry measurements. In J. Nývlt and S. Žáček, editors, *Industrial Crystallization 87*, pages 415–418, Bechyně, Czechoslovakia, 1987. Elsevier Science Publishers.

[75] J. Jager, S. de Wolf, H. J. M. Kramer, and E. J. de Jong. Estimation of nucleation kinetics from crystal size distribution transients of a continuous crystallizer. *Chemical Engineering Science*, 46(3):807–818, 1991.

[76] Johan Jager, Sjoerd de Wolf, Herman J. M. Kramer, and Esso J. de Jong. Estimation of nucleation and attrition kinetics from csd transients in a continuous crystallizer. *Transactions of the Institute of Chemical Engineers*, 69:53–62, 1991.

[77] A. H. Janse and E. J. de Jong. The occurence of growth dispersion and its consequences. In J. W. Mullin, editor, *Industrial Crystallization*, New York, 1976. Plenum Press.

[78] Harold Jeffreys. *Theory of Probability.* Oxford University Press, London, 3rd edition, 1961.

[79] R. T. Johnson, R. W. Rousseau, and W. L. McCabe. Factors affecting contact nucleation. *AIChE Symposium Series*, 68(121):31–41, 1972.

[80] A. G. Jones. Optimal operation of a batch cooling crystallizer. *Chemical Engineering Science*, 29:1075–1087, 1974.

[81] A. G. Jones, J. Budz, and J. W. Mullin. Crystallization kinetics of potassium sulfate in an MSMPR agitated vessel. *AIChE Journal*, 32(12):2002–2008, 1986.

[82] A. G. Jones, J. Budz, and J. W. Mullin. Batch crystallization and solid-liquid separation of potassium sulphate. *Chemical Engineering Science*, 42(4):619–629, 1987.

[83] A. G. Jones and J. Mydlarz. Slurry filtrability of potash alum. *The Canadian Journal of Chemical Engineering*, 68(3):513–518, 1990.

[84] Noriaki Kubota and Masato Fujiwara. Minimum seed crystal size for secondary nucleation of potassium alun in a stirred-vessel crystallizer. *Journal of Chemical Engineering of Japan*, 23(6):691–696, 1990.

[85] H. W. Lee and E. G. Saleeby. Mathematical behavior of the population balance of a msmprc with agglomeration, feed and size-dependent growth rate. *Computers & Chemical Engineering*, 18(10):899–907, 1994.

[86] B. M. Liang, R. W. Hartel, and K. A. Berglund. Contact nucleation in sucrose crystallization. *Chemical Engineering Science*, 42(11):2723–2727, 1987.

[87] B. M. Liang, R. W. Hartel, and K. A. Berglund. Growth rate dispersion in seeded batch sucrose crystallization. *AIChE Journal*, 33(12):2077–2079, 1987.

[88] I. Liszi, S. Halász, and B. Bodor. Method for determination of supersaturation using conductivity for salting-out crystallization. In J. Nývlt and S. Žáček, editors, *Industrial Crystallization 87*, pages 399–402, Amsterdam, 1987. Elsevier Science Publishers.

[89] J. D. Litster, D. J. Smit, and M. J. Hounslow. Adjustable discretized population balance for growth and agglomeration. *AIChE Journal*, 41(3):591–603, March 1995.

[90] I. Livk, M. Gregorka, and C. Pohar. Identification of batch crystallization processes (estimate of kinetic parameters of sodium-perborate precipitation). *Computers & Chemical Engineering*, 19:S241–S246, 1995.

[91] Dimitris P. Lymberopoulos and Jay D. Schieber. Stochastic dynamic simulation of the boltzmann equation for electron swarms in glow discharges. *Physical Review E*, 50(6):4911–4919, 1994.

[92] P. Marchal, R. David, J. P. Klein, and J. Villermaux. Crystallization and precipitation engineering—1. An efficient method for solving population balance in crystallization with agglomeration. *Chemical Engineering Science*, 43(1):59–67, 1988.

[93] H. B. Matthews, Stephen M. Miller, and James B. Rawlings. Model identification for crystallization: Theory and experimental verification. *Powder Technology*, 88:227–235, 1996.

[94] S. M. Miller and J. B. Rawlings. Model identification and control strategies for batch cooling crystallizers. *AIChE Journal*, 40(8):1312–1327, August 1994.

[95] Stephen M. Miller. *Modelling and Quality Control Strategies for Batch Cooling Crystallizers*. PhD thesis, The University of Texas at Austin, 1993.

[96] Manfred Morari. Some comments on the optimal operation of batch crystallizers. *Chemical Engineering Communications*, 4:167–171, 1980.

[97] J. W. Mullin. *Crystallization*. Butterworth & Co (Publishers) Ltd, London, England, 3rd edition, 1993.

[98] J. W. Mullin and C. J. Leci. Desupersaturation of seeded citric acid solutions in a stirred vessel. *AIChE Symposium Series*, 68(121):8–20, 1972.

[99] J. W. Mullin and J. Nývlt. Programmed cooling of batch crystallizers. *Chemical Engineering Science*, 26:369–377, 1971.

[100] Larry E. Murray and E. K. Reiff, Jr. Design of transient experiments for identification of fixed bed thermal transport properties. *The Canadian Journal of Chemical Engineering*, 62:55–61, February 1984.

[101] J. Mydlarz and A. G. Jones. Solubility and density isotherms for potash alum-water-acetone. *Journal of Chemical and Engineering Data*, 34:365–367, 1989.

[102] J. Mydlarz and A. G. Jones. On the estimation of size-dependent crystal growth rate functions in msmpr crystallizers. *The Chemical Engineering Journal*, 53:125–135, 1993.

[103] J. Nývlt, M. Karel, and S. Pisăřík. Measurement of supersaturation. *Crystal Research and Technology*, 29:409–415, 1994.

[104] E. P. K. Ottens, A. H. Janse, and E. J. De Jong. Secondary nucleation in a stirred vessel cooling crystallizer. *Journal of Crystal Growth*, 13/14:500–505, 1972.

[105] S. Ou and J. Seader. Contribution of multiple scattering to light transmittance by a collimated beam. *Computers & Chemical Engineering*, 11(6):757–761, 1987.

[106] B. Palwe, M. Chivate, and N. Tavare. Growth kinetics of ammonium nitrate crystals in a draft tube baffled batch crystallizer. *Industrial and Engineering Chemistry Process Design and Development*, 24:914–919, 1985.

[107] Robert H. Perry, Don W. Green, and James O. Maloney, editors. *Perry's Chemical Engineer's Handbook*. McGraw-Hill Book Company, New York, 6th edition, 1984.

[108] L. R. Petzold. A description of DASSL: A differential-algebraic system solver. In R. S. Stepleman, editor, *Scientific Computing*, pages 65–68, Amsterdam, 1983. North-Holland.

[109] J. C. Pinto, M. W. Lobão, and J. L. Monteiro. Sequential experimental design for parameter estimation: Analysis of relative deviations. *Chemical Engineering Science*, 46(12):3129–3138, 1991.

[110] D. J. Pritchard and D. W. Bacon. Prospects for reducing correlations among parameter estimates in kinetic models. *Chemical Engineering Science*, 33:1539–1543, 1978.

[111] Yanfeng Qiu and Åke C. Rasmuson. Growth and dissolution of succinic acid crystals in a batch stirred crystallizer. *AIChE Journal*, 36(5):665–676, 1990.

[112] Yanfeng Qiu and Åke C. Rasmuson. Nucleation and growth of succinic acid in a batch cooling crystallizer. *AIChE Journal*, 37(9):1293–1304, 1991.

[113] Yanfeng Qiu and Åke C. Rasmuson. Estimation of crystallization kinetics from batch cooling experiments. *AIChE Journal*, 40(5):799–812, 1994.

[114] D. Ramkrishna. Statistical models of cell populations. In T. K. Ghose, A. Fiechter, and N. Blakebrough, editors, *Advances in Biochemical Engineering*, volume 11, page 1, 1979.

[115] D. Ramkrishna. Analysis of population balance—IV: The precise connection between monte carlo simulation and population balances. *Chemical Engineering Science*, 36:1203–1209, 1981.

[116] D. Ramkrishna. The status of population balances. *Reviews in Chemical Engineering*, 3(1):49–95, 1985.

[117] D. Ramkrishna and J. D. Borwanker. A puristic analysis of population balance–I. *Chemical Engineering Science*, 28:1423–1435, 1973.

[118] D. Ramkrishna and J. D. Borwanker. A puristic analysis of population balance–II. *Chemical Engineering Science*, 29:1711–1721, 1974.

[119] D. Ramkrishna, B. H. Shah, and J. D. Borwanker. Analysis of population balance–iii. *Chemical Engineering Science*, 31:435–442, 1976.

[120] A. D. Randolph and M. A. Larson. Transient and steady-state size distributions in continuous mixed suspension crystallizers. *AIChE Journal*, 8(5):639–645, 1962.

[121] A. D. Randolph and E. T. White. Modeling size dispersion in the prediction of crystal-size distribution. *Chemical Engineering Science*, 32:1067–1076, 1977.

[122] Alan D. Randolph and Michael D. Cise. Nucleation kinetics of the potassium sulfate-water system. *AIChE Journal*, 18(4):798–807, July 1972.

[123] Alan D. Randolph and Maurice A. Larson. *Theory of Particulate Processes*. Academic Press, San Diego, second edition, 1988.

[124] J. B. Rawlings, S. M. Miller, and W. R. Witkowski. Model identification and control of solution crystallization processes: A review. *Industrial and Engineering Chemistry Research*, 32(7):1275–1296, July 1993.

[125] J. B. Rawlings and W. H. Ray. The modelling of batch and continuous emulsion polymerization reactors. Part I: Model formulation and sensitivity to parameters. *Polymer Engineering & Science*, 28:237–256, 1988.

[126] J. B. Rawlings and W. H. Ray. The modelling of batch and continuous emulsion polymerization reactors. Part II: Comparison with experimental data from continuous stirred tank reactors. *Polymer Engineering & Science*, 28:257–273, 1988.

[127] Terry P. Redman and Sohrab Rohani. On-line determination of supersaturation of a kcl-nacl aqueous solution based on density measurement. *The Canadian Journal of Chemical Engineering*, 72:64–71, February 1994.

[128] Park M. Reilly, R. Bajramovic, G. E. Blau, D. R. Branson, and M. W. Sauerhoff. Guidelines for the optimal design of experiments to estimate parameters in first order kinetic models. *The Canadian Journal of Chemical Engineering*, 55:614–622, October 1977.

[129] U. Riebel, V. Kofler, and F. Löffler. Control of supersaturation in instationary suspension crystallization. In A. Mersmann, editor, *Industrial Crystallization 90*, pages 595–599, Amsterdam, 1990.

[130] Sohrab Rohani, N. S. Tavare, and John Garside. Control of crystal size distribution in a batch cooling crystallizer. *The Canadian Journal of Chemical Engineering*, 68:260–267, April 1990.

[131] R. W. Rousseau, K. K. Li, and W. L. McCabe. The influence of seed crystal size on nucleation rates. *AIChE Symposium Series*, 72(153):48–52, 1976.

[132] D. F. Rudd. A generalization of the residence time concept. *The Canadian Journal of Chemical Engineering*, pages 197–202, 1962.

[133] E. G. Saleeby and H. W. Lee. Solution and analysis for crystallization with agglomeration. *Chemical Engineering Science*, 49(12):1879–1884, 1994.

[134] E. G. Saleeby and H. W. Lee. On the solution of the PBE with agglomeration and random growth rate dispersion. *Chemical Engineering Science*, 50(12):1971–1981, 1995.

[135] R. W. H. Sargent and G. R. Sullivan. The development of an efficient optimal control package. In J. Stoer, editor, *Lecture Notes in Control and Information Sciences 7*, pages 158–168, Berlin, 1978. 8th IFIP Conference on Optimization Techniques, Springer-Verlag.

[136] B. H. Shah, D. Ramkrishna, and J. D. Borwanker. Simulation of bubble populations in a gas fluidized bed. *Chemical Engineering Science*, 32:1419–1425, 1977.

[137] B. H. Shah, D. Ramkrishna, and J. D. Borwanker. Simulation of particulate systems using the concept of the interval of quiescence. *AIChE Journal*, 23(6):897–904, 1977.

[138] M. B. Sherwin, R. Shinnar, and S. Katz. Dynamic behavior of the well-mixed isothermal crystallizer. *AIChE Journal*, 13:1141–1153, 1967.

[139] Subhas K. Sikdar and Alan D. Randolph. Secondary nucleation of two fast growth systems in a mixed suspension crystallizer: Magnesium sulfate and citric acid water systems. *AIChE Journal*, 22(1):110–117, 1976.

[140] P. N. Singh and D. Ramkrishna. Solution of population balance equations by MWR. *Computers & Chemical Engineering*, 1:23–31, 1977.

[141] W. E. Stewart, M. Caracotsios, and J. P. Sørensen. *Computer-Aided Modelling of Reactive Systems*. 1996. In preparation.

[142] W. E. Stewart, T. L. Henson, and G. E. P. Box. Model discrimination and criticism with single-response data. *AIChE Journal*, 42(11):3055–3062, 1996.

[143] Thomas L. Sutton and John F. MacGregor. The analysis and design of binary vapour-liquid equilibrium experiments; part ii: The design of experiments. *The Canadian Journal of Chemical Engineering*, 55:609–613, October 1977.

[144] N. S. Tavare. Batch crystallizers: A review. *Chemical Engineering Communications*, 61:259–318, 1987.

[145] F. M. Tiller and J. R. Crump. Solid-liquid separation: An overview. *Chemical Engineering Progress*, 73(11):65–75, October 1977.

[146] D. C. Timm and M. A. Larson. Effect of nucleation kinetics on the dynamic behavior of a continuous crystallizer. *AIChE Journal*, 14:452–457, 1968.

[147] Michael Ulrich. Optimization of batch solution crystallization. *German Chemical Engineering*, 4:195–200, 1979.

[148] H.C. van de Hulst. *Light Scattering by Small Particles*. Dover Publications, New York, 1981.

[149] A. E. D. M. van der Heijden, J. P. van der Eerden, and G. M. van Rosmalen. The secondary nucleation rate: A physical model. *Chemical Engineering Science*, 49(18):3103–3113, 1994.

[150] N. G. van Kampen. *Stochastic Processes in Physics and Chemistry.* Elsevier Science Publishers, Amsterdam, The Netherlands, 2nd edition, 1992.

[151] J. V. Villadsen and W. E. Stewart. Solution of boundary-value problems by orthogonal collocation. *Chemical Engineering Science*, 22:1483–1501, 1967.

[152] H. Wakao, H. Hiraguchi, and T. Ishii. A simulation of crystallization from aqueous supersaturated solutions in a batch isothermal stirred tank. *The Chemical Engineering Journal*, 35(7):169–178, 1987.

[153] Abraham Wald. On the efficient design of statistical investigations. *Annals of Mathematical Statistics*, 14:134–140, 1943.

[154] A. G. Walton. Nucleation in liquids and solutions. In A. C. Zettlemoyer, editor, *Nucleation*, New York, 1969. Marcel Dekker.

[155] Maw-Ling Wang, Huei-Neng Horng, and Rong-Yeu Chang. Modeling of crystallization systems via generalized orthogonal polynomials method. *Chemical Engineering Communications*, 57:197–213, 1987.

[156] S. Wang and A. Mersmann. Initial-size-dependent growth rate dispersion of attrition fragments and secondary nuclei. *Chemical Engineering Science*, 47(6):1365–1371, 1992.

[157] W. R. Witkowski and J. B. Rawlings. Modelling and control of crystallizers. In *Proceedings of the 1987 American Control Conference*, pages 1400–1405, Minneapolis, Minnesota, June 1987. American Control Conference.

[158] Walter R. Witkowski. *Model Identification and Parameter Estimation of Crystallization Processes.* PhD thesis, The University of Texas at Austin, 1990.

[159] Walter R. Witkowski, Stephen M. Miller, and James B. Rawlings. Light scattering measurements to estimate kinetic parameters of crystallization. In Allan S. Myerson and Ken Toyokura, editors, *Crystallization as a Separations Process*, pages 102–114, Washington, DC, 1990. American Chemical Society.

[160] R. C. Zumstein and R. W. Rousseau. Growth rate dispersion by inital growth rate distributions and growth rate fluctuations. *AIChE Journal*, 33(1):121–129, 1987.

Vita

Hazel Benton Matthews III (a.k.a. Rock) is the son of Hazel Benton Matthews Jr. (a.k.a. Skip) and Betty Matthews. He was born in Madison, Wisconsin on May 17, 1968 as fitting celebration of the completion of his father's Ph.D. While a toddler, he lived briefly in San Francisco, California before returning to the Matthews' ancestral homeland in the state of North Carolina. He graduated from C. E. Jordan High School, Durham, NC in 1986 and in the Fall of that year attended North Carolina State University. He received his Bachelor of Science degree in Chemical Engineering in May, 1990 from North Carolina State University. His graduate career began in the Fall of 1990 when he joined the research group of Dr. James B. Rawlings in the Chemical Engineering Department at The University of Texas at Austin. In June 1995, he moved with Dr. Rawlings from Austin to the Chemical Engineering Department at the University of Wisconsin-Madison.

Permanent Address: 4-B West Atlanta
Wrightsville Beach, NC 28480

This dissertation was prepared with LaTeX 2_ε[1] by the author.

[1]This particular University of Wisconsin complient style was carved from The University of Texas at Austin styles as written by Dinesh Das (LaTeX 2_ε), Khe–Sing The (LaTeX), and John Eaton (LaTeX). Knives and chisels wielded by John Campbell and Rock Matthews.

Index