MODELLING, SIMULATION, AND CONTROL OF POLYMORPHIC CRYSTALLIZATION

MARTIN WIJAYA HERMANTO

(*B. Eng. (Hons.), NUS*)

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF CHEMICAL AND BIOMOLECULAR ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2008

Acknowledgements

I would like to express my deep and sincere gratitude to my supervisors, Dr. Min-Sen Chiu and Dr. Richard D. Braatz, for their constant support and guidance throughout my PhD study at National University of Singapore. This thesis would not have been possible without their inspiration, encouragement and their detailed and constructive comments on my research.

I sincerely thank Dr. Lakshminarayanan Samavedham and Dr. Qing-Guo Wang for their priceless inputs and advices during PhD oral examination. I warmly thank all my lab mates, Dr. Yasuki Kansha, Ye Myint Hlaing, Ankush Gameshreddy Kalmukale, Bu Xu, Yan Li, Xin Yang, and Imma Nuella for their moral support and help rendered to me. My special thanks to Mr. Boey Kok Hong who has provided his continuous assistance to make a workstation ready and available for my research and to all academic and administrative staffs in the Chemical and Biomolecular Engineering Department who have directly and indirectly help with my research work. I am indebted to the National University of Singapore for the outstanding research facilities and the research scholarship provided for my research.

I am very grateful to my wife, Fonny, for her constant moral supports, encouragements, and prayers which fire up my spirit when I faced obstacles and difficulties in my research. I am greatly indebted to my parents and family for their concerns and supports. Last but not least, I thank and praise my God, Jesus Christ, for His grace and wisdom, without which I will not be able to finish this research.

Contents

Summary

Polymorphism, in which multiple crystal forms exist for the same chemical compound, is of significant interest to industry. The variation in physical properties such as crystal shape, solubility, hardness, colour, melting point, and chemical reactivity makes polymorphism an important issue for the food, specialty chemical, and pharmaceutical industries, where products are specified not only by chemical composition, but also by their performance. Controlling polymorphism to ensure consistent production of the desired polymorph is important in those industries, including drug manufacturing where safety is paramount. In this thesis, the modelling, simulation, and control of polymorphic crystallization of L-glutamic acid, comprising the metastable α-form and the stable β-form crystals, are investigated.

With the ultimate goal being to better understand the effects of process conditions on crystal quality and to control the formation of the desired polymorph, a kinetic model for polymorphic crystallization of L-glutamic acid based on population balance equations is developed using Bayesian inference. Such a process model can facilitate the determination of optimal operating conditions and speed process

development, compared to time-consuming and expensive trial-and-error methods for determining the operating conditions. The developed kinetic model appears to be the first to include all of the transformation kinetic parameters including dependence on the temperature, compared to past studies on the modelling of L-Glutamic acid crystallization.

Next, numerical simulation of the developed model is investigated. It is important to have efficient and sufficiently accurate computational methods for simulating the population balance equations to ensure the behaviour of the numerical solution is determined by the assumed physical principles and not by the chosen numerical method. In this thesis, the high-order weighted essentially non-oscillatory (WENO) methods are investigated and shown to give better computational efficiency compared to the high resolution (HR) and the standard second-order finite difference (FD2) methods to simulate the model of polymorphic crystallization of L-glutamic acid developed in this thesis.

In non-polymorphic crystallization, the two most popular control strategies are the temperature control (T-control) and concentration control (C-control) strategies. In this study, the robustness of these control strategies are investigated in polymorphic crystallization using the model developed in this thesis. Simulation studies show that T-control is not robust to kinetics perturbations, while C-control performs very robustly but long batch times may be required.

Despite the high impact of model predictive control (MPC) in academic research and industrial practice, its application to solution crystallization processes has been

rather limited and there is no published result on the implementation of MPC or nonlinear MPC (NMPC) to a polymorphic crystallization, which is more challenging. In this thesis, an efficient NMPC strategy based on the extended predictive self-adaptive control (EPSAC) which does not rely on nonlinear programming is developed for the polymorphic transformation process. Compared to the T-control, C-control, and quadratic matrix control with successive linearization (SL-QDMC), simulation results show that the NMPC strategy gives good overall robustness while satisfying all constraints on manipulated and state variables within the specified batch time.

Finally, exploiting the repetitive nature of batch processes, an integrated nonlinear model predictive control and batch-to-batch (NMPC-B2B) control strategy based on a hybrid model is developed for the polymorphic transformation process. The hybrid model consists of a first-principles model and a PLS model, where information from the previous batches are utilized to update the control trajectory in the next batch. The proposed NMPC-B2B strategy allows the NMPC to perform online control which handle the constraints effectively while the batch-to-batch control refines the model by learning from the previous batches. Compared to the standard batch-to-batch (B2B) control strategy, the proposed NMPC-B2B control strategy gives better performance where it satisfies all the state constraints and produces faster and smoother convergence. In addition, it is verified that through the learning process, both B2B and NMPC-B2B control strategies are more advantageous to be employed to address the plant-model mismatch in an effective manner.

List of Tables

List of Figures

Nomenclature

$a_{i,1}$, $a_{i,2}$, $a_{i,3}$	Parameters for the saturation concentration of the i-form crystals
$\bar{a}_{k+1/2}$	Roe speed
B_i	Nucleation rate of the i-form crystals
C	Solute concentration
$C_{\mathrm{sat},i}$	Saturation concentration of the i-form crystals
d_k	Unmeasured disturbances
$\delta(\cdot)$	Dirac delta function
E, F	Residuals matrices
E_{gi}	Activation energy for the growth rate of i-form crystals
E_{L_1}	Prediction errors in terms of the L_1 norm
$E[\cdot]$	Expected value
ϵ	Vector of slack variables
f	System dynamics function

f_i, $f_{\text{seed},i}$, $f_{\text{nucl},i}$	Total, seed, and nucleated crystal size distribution of the i-form crystals
G_i	Growth rate of the i-form crystals
g	Measurement function
g_j	The jth step response coefficient
h	Linear and nonlinear constraints for the system
h_j	The jth impulse response coefficient
I_k	The kth cell
IS_m	Smoothness indicator
K_k	Kalman gain
$k_{b\alpha}$, $k_{g\alpha}$, $k_{d\alpha}$	Nucleation, growth, and dissolution rates of α-form crystals
$k_{b\beta,j}$, $k_{g\beta,j}$	The jth nucleation and growth rates of β-form crystals
$k_{gi,0}$	Pre-exponential factor for the growth rate of i-form crystals
k_{vi}	Volumetric shape factor of the i-form crystals
L, L_0	Characteristic length of crystals and nuclei
$\mathrm{L}(\theta \mid \mathbf{y})$	Likelihood of θ
L_k	Characteristic length of crystals at the kth discretized point

ΔL	Discretization size of crystal length
λ_i **(**κ_i**)**	Scaling factor for the seed crystal size distribution of i-form crystals
$\mu_{\beta,3}^{\text{nucl}}$	The third moment of the nucleated β-form crystals
$\mu_{\beta,3}^{\text{seed}}$	The third moment of the seed of β-form crystals
$\mu_{i,n}$	The nth moment of the i-form crystals
$\mu_{\text{seed},i}$	Mean for the seed crystal size distribution of i-form crystals
N	Total samples in a batch
N_{d_j}	Number of time samples of j^{th} variable in Bayesian inference
N_m	Number of measured variables in Bayesian inference
N_s	Total number of values drawn from the second halves for all the chains
n_{latent}	Number of the latent variables used in PLS
n_u	Number of inputs
n_x	Number of system states
n_y	Number of measured variables
O,Q	Matrices of loadings for $\mathbf{X}$ and $\mathbf{Y}$
O_{L_1}	L_1 self convergence order

ω_m	Scalar weight to each candidate stencil S_m for the flux approximation
P, P_d	Predicted and desired final product quality
P_1	The first product quality: mass of β-form crystals
P_2	The second product quality: ratio of the nucleated crystal mass to the seed crystal mass of β-form crystals
$\Pr$	Probability function
$\Pr(\theta)$	Prior distribution of θ
$\Pr(\mathbf{y} \mid \theta)$	Sampling distribution (or data distribution) for fixed parameters θ
$P_{x_{a,k-1}}$	Augmented states covariance at the previous sampling time
$P_{x_a y}$	Predicted cross-covariance matrix between the augmented system states and the measured variables
$\hat{p}_{k+1/2}$, $\hat{p}_{k-1/2}$	Numerical flux approximation at node $k+1/2$ and $k-1/2$
ϕ	Flux limiter
q_m^r	Quadratic polynomial flux approximation function

$\hat{R}_i$	Potential scale reduction factors
ρ_i	Density of the i-form crystals
ρ_{solv}	Density of the solvent
S, U	Matrices of scores for $\mathbf{X}$ and $\mathbf{Y}$
S_i	Supersaturation of the i-form crystals
S_m	Candidate stencil
σ_j	Standard deviation of the measurement noise in the j^{th} variable
$\sigma_{\text{seed},i}$	Standard deviation for the seed crystal size distribution of i-form crystals
$\chi_{i,k-1}$	Scaled sigma points
$\chi^{\circ}_{i,k-1}$	Unscaled sigma points
T	Crystallizer temperature
T_{min}**,** T_{max}	Minimum and maximum temperatures due to the limitation of water bath heating/cooling
Θ	Approximated samples from the target distribution
θ	A vector of unknown parameters of interest
$\theta_i^{c,s}$	Simulation draws of parameter i from step chain c at step s
θ_{min}**,** θ_{max}	Minimum and maximum values of θ

$u_{b,k+i}$	Predetermined future control scenario
u_k, u_k^j	Process inputs
δu_{k+i}	Optimizing future control actions
v_k	Noise on the measured variables
$\mathbf{W}_{\Delta\mathcal{U}}$, $\mathbf{W}_{\mathrm{d}\mathcal{U}}$	Weight matrices which penalize excessive changes in the input variable which occur within-batch and inter-batch in the B2B control strategy
$\mathbf{W}_{\Delta\mathbf{U}}$, $\mathbf{W}_{\mathrm{d}\mathbf{U}}$	Weight matrices which penalize excessive changes in the input variable which occur within-batch and inter-batch in the NMPC-B2B control strategy
$\mathbf{W}_p$	Weight matrix for the product quality in NMPC strategy
$\mathbf{W}_u$	Weight matrix for the change in input variables in NMPC strategy
W_p	Scalar weight corresponding to the final product quality in B2B/NMPC-B2B control strategy
w_k	Noise on the system states
X, Y	Database matrices for PLS model
$x_{a,k}$	Augmented system states

ξ_k	Noise on the unmeasured disturbances
x_k	System states at the kth sampling instance
$\mathbf{y}$	Collected data which is used to infer θ
y_{jk}, $\hat{y}_{jk}$	Measurement and predicted value of j^{th} variable at sampling instance k, respectively
y_k	Measured variables
$z_{b,k+i}$	Part of z_{k+i} calculated using the nonlinear model and predetermined sequence $u_{b,k+i}$
$z^j_{\text{fp},k}$	Part of z^j_k calculated using the first-principles model with nominal model parameters
z_{k+i}	Process variable of interest
z^j_k	Process variable of interest at sampling instant k and batch j
$z_{l,k+i}$	Part of z_{k+i} calculated using impulse response coefficient
$\Delta z^j_{\text{pls},k}$	Part of z^j_k calculated using the PLS model

Abbreviations

ATR-FTIR	Attenuated total reflection Fourier transform infrared
B2B	Batch-to-batch
BMPC	Batch model predictive control
C-control	Concentration control
CLD	Chord length distribution
CSD	Crystal size distribution
DE	Differential evolution
DE-MC	Differential evolution Markov chain
EKF	Extended Kalman filter
ENO	Essentially non-oscillatory
EPSAC	Extended predictive self-adaptive control
FBRM	Focused beam reflectance measurement
FD2	Second-order finite difference
FVM	Finite volume method

HR	High resolution
ILC	Iterative learning control
JSHWENO	Jiang and Shu's version of WENO with Henrick mapping
KF	Kalman filter
LOCWENO	Liu et al's version of WENO
LTV	Linear time varying
MCMC	Markov Chain Monte Carlo
MPC	Model predictive control
NMPC	Nonlinear model predictive control
NMPC-B2B	Integrated nonlinear model predictive control and batch-to-batch
ODE	Ordinary differential equation
PBE	Population balance equation
PBM	Population balance model
PCR	Principal component regression
PDE	Partial differential equation
PLS	Partial least squares
PSD	Particle size distribution
Q-ILC	Quadratic criterion-based iterative learning control

QDMC	Quadractic dynamic matrix control
QP	Quadratic programming
QPLS	Quadratic partial least squares
SL-QDMC	Quadractic dynamic matrix control with successive linearization
T-control	Temperature control
TVD	Total diminishing variation
UKF	Unscented Kalman filter
UT	Unscented transformation
WENO	Weighted essentially non-oscillatory
Wpower-ENO	Weighted power ENO method
XRD	X-ray diffraction

Chapter 1

Introduction

1.1 Motivation

Crystallization is one of the oldest unit operations and remains the most utilized purification technique in pharmaceutical industries. In fact, most pharmaceutical manufacturing processes include a series of crystallization processes where their product quality is often associated with the crystal final form (such as crystal habit, shape and size distribution). Unfortunately, despite its long history, crystallization process is still not very well understood as it involves many complex mechanisms (e.g., fine dissolution, agglomeration, growth dispersion, etc) in addition to the main ones (i.e. nucleation and growth). This makes controlling crystallization process very challenging.

Recently, there is a rapid growth of interest in polymorphism [6, 30, 42, 132, 172]. It is a phenomenon that a substance can have more than one crystal form. This

phenomenon was first observed in 1798 by Klaproth. He discovered the calcium carbonate polymorphs, namely, calcite and aragonite. In 1899, Ostwald concluded that almost every substance can exist in two or more solid phases provided the experimental conditions are suitable. According to Ostwald's Rule of Stages, in a polymorphic system, the most soluble metastable form will always appear first, followed by the more stable one.

The appearance of metastable phases is associated with the environmental conditions at the time of precipitation and as a result it is of considerable importance in biomineralization, diagenesis and synthetic industrial chemistry. In the latter context, metastable solid phases are commonly encountered in the production of specialty chemicals such as pharmaceuticals, dyestuffs and pesticides. Deliberate isolation of these phases is sometimes effected because of their advantageous processing or application properties. In other cases, however, the formulation of a product as a metastable phase may be unacceptable because of subsequent phase transformation and crystal growth, which could occur during storage and result in product degradation [21].

Morris et al. [111] stated that unexpected or undesired polymorphic transformation of pharmaceutical is not uncommon during manufacturing processes including crystallization process. For example, in 1998, Abbot laboratories withdrew its HIV drug, Ritonavir, because of the unexpected appearance of a new crystal form that had different dissolution and absorption characteristics from the standard product. The two crystal forms, which have the same molecular structure, are distinguished

by the way in which the molecules are packed within the crystals, and each has distinct physical and thermodynamic properties [6]. In addition, to highlight the importance of polymorphism in the pharmaceutical industry, the U.S. Food and Drug Administration (FDA) has tightened regulations for new drug applications to ensure that the drugs contain only the desired polymorph.

The variations in physical properties such as crystal shape, solubility, hardness, colour, melting point, and chemical reactivity make polymorphism an important issue for the food, speciality chemical and pharmaceutical industries, where products are specified not by chemical composition only, but also by their performance [6]. As a result, controlling polymorphism to ensure consistent production of the desired polymorph is very crucial in those industries, including drug manufacturing industry where safety is of paramount importance.

Encouraged by the importance of polymorphism in pharmaceutical industries, this study investigates the modelling, simulation, and control of polymorphic crystallization of L-glutamic acid, which consists of the metastable α-form and the stable β-form crystals.

1.2 Contributions

The main contributions of this thesis in the area of modelling, simulation, and control of polymorphic crystallization process can be summarized as follows:

(1) Process model can facilitate the determination of optimal operating condi-

tions and speed up process development in pharmaceutical industries. In this study, a kinetic model of L-glutamic acid polymorphic crystallization is developed from batch experiments with in-situ measurements including attenuated total reflection Fourier transform infrared (ATR-FTIR) spectroscopy which is used to infer the solute concentration and focused beam reflectance measurement (FBRM) which provides crystal size information. Bayesian inference is employed to obtain the posterior distribution for the model parameters, which can be used to quantify the accuracy of model predictions and can be incorporated into robust control strategies for crystallization process [114]. Furthermore, the developed kinetic model appears to be the first to include all of the transformation kinetic parameters including dependence on the temperature, compared to past studies on the modelling of L-Glutamic acid crystallization [115, 139].

(2) Numerical simulation of the developed model is important in the investigation of the effects of various operating conditions on the polymorphic crystallization and can be used for optimal design and control [64, 130, 139]. Therefore, it is indispensable to select an efficient and sufficiently accurate computational method for simulating the model to ensure the behaviour of the numerical solution is determined by the assumed physical principles and not by the chosen numerical method. In this study, high-order numerical simulation techniques based on the weighted essentially non-oscillatory (WENO)

methods are investigated and shown to give better computational efficiency compared to the high resolution (HR) finite volume method and a second-order finite difference (FD2) method to simulate the model of polymorphic crystallization of L-glutamic acid developed in this thesis.

(3) The two most popular control strategies implemented in non-polymorphic crystallization processes have been the temperature control (T-control) and concentration control (C-control) strategies. In this study, these control strategies are implemented in the polymorphic transformation process using the model developed in this thesis. Simulation studies show that T-control is not robust to kinetics perturbations, while C-control performs very robustly but long batch times may be required.

(4) Model predictive control (MPC) strategy is widely recognized as a powerful technique to address industrially important control problems. However, its implementation to crystallization processes has been rather limited [35, 79, 113, 131, 155] and there is no published result on the implementation of MPC or nonlinear MPC (NMPC) to a polymorphic crystallization, which is more challenging for a number of reasons. First, the phase equilibria and crystallization kinetics are more complicated. Second, the method of moments heavily used in past control algorithms for crystallization processes does not apply during a polymorphic transformation, so that the full PDEs need to be solved. As a consequence, the computation time required increases

considerably which prohibits the straightforward application of nonlinear programming. In this study, a practical NMPC strategy based on extended predictive self-adaptive control (EPSAC) [32, 34, 70, 134, 156] is developed for the polymorphic transformation of L-glutamic acid from the metastable α-form to the stable β-form. To implement the proposed NMPC strategy, an unscented Kalman filter (UKF) [74–78] is utilized to estimate the unmeasurable states. Compared to the T-control, C-control, and quadratic matrix control with successive linearization (SL-QDMC), the NMPC strategy shows good overall robustness while satisfying all constraints on manipulated and state variables within the specified batch time.

(5) Exploiting the fact that batch processes are repetitive in nature, it is possible to implement batch-to-batch (B2B) control to the polymorphic crystallization process considered in this study, which uses information from previous batches to update the process model in order to iteratively compute the optimal operating conditions for each batch. However, due to the open-loop nature of batch-to-batch control, this optimal policy is not implemented until the next batch. As a result, when the process model is still not accurate, which is likely the case in the first few batches, it is possible that the input or/and output constraints will be violated. Therefore, in this study we propose an integrated nonlinear model predictive control and batch-to-batch (NMPC-B2B) control strategy based on a hybrid model. The hybrid model compris-

ing the nominal first-principles model and a correction factor based on an updated partial least square (PLS) model is utilized to predict the process variables and final product quality. In the proposed NMPC-B2B control strategy, the NMPC performs online control to handle the constraints effectively while the batch-to-batch control refines the model by learning from the previous batches. Simulation studies show that the proposed NMPC-B2B control strategy produces faster and smoother convergence and satisfies all the state constraints, compared to the standard B2B control strategy. Furthermore, the learning process in both B2B and NMPC-B2B control strategies counteracts the plant-model mismatch effectively after several batches.

1.3 Thesis Organization

This thesis is organized as follows. In the next chapter, literature review on the fundamental of crystallization and the recent development of the modelling, simulation, and control of crystallization process is presented. Chapter 3 presents the modelling of the L-glutamic acid polymorphic crystallization, followed by the investigation on the high-order simulation of polymorphic crystallization in Chapter 4. The control strategies which includes the temperature control (T-control), concentration control (C-control), nonlinear model predictive control (NMPC), and batch-to-batch control strategies are discussed in Chapters 5 to 7. Finally, conclusions from the present work and suggestions for the future work are given in Chapter 8.

Chapter 2

Literature Review

This chapter discusses the fundamental of crystallization which includes the definition, driving force, mechanism, and polymorphism. Subsequently, the recent development on the modelling, simulation and control of crystallization is reviewed.

2.1 Crystallization fundamentals

Crystallization is a supramolecular process by which an ensemble of randomly organized molecules, ions or atoms in a fluid come together to form an ordered three-dimensional molecular array which is called crystal [29]. Crystallization is indispensable in drug manufacturing as it is the main separation and purification process. Not only does crystallization affect the efficiency of downstream operations such as filtering and drying, the efficacy of the drug can be dependent on the final crystal form [42].

To understand crystallization process, it is important to know the key elements of crystallization which are discussed next.

2.1.1 The driving force for crystallization

As with any chemical rate process, crystallization is a kinetic process which is driven by concentration. However, in crystallization, the concentration range over which the process can occur is limited by the equilibrium composition of the system corresponding to the conditions chosen [29]. Figure 2.1 shows a hypothetical solubility curve. A solution whose composition lies below the solubility curve is undersaturated and existing crystals will dissolve. A solution lying above solubility curve is termed supersaturated, since the amount of dissolved solute is greater than the equilibrium saturation value. Crystals can nucleate and grow only if the solution is supersaturated and so the production of a supersaturated solution is a prerequisite for crystallization. Supersaturation is typically created by cooling, evaporation, and/or addition of antisolvent, including changing the pH by addition of acid or base.

There is a region above solubility curve called metastable zone. In this zone, though existing crystals will grow, it is difficult to create new crystals. Once this zone is exceeded, new crystals form spontaneously and the solution is now labile [29].

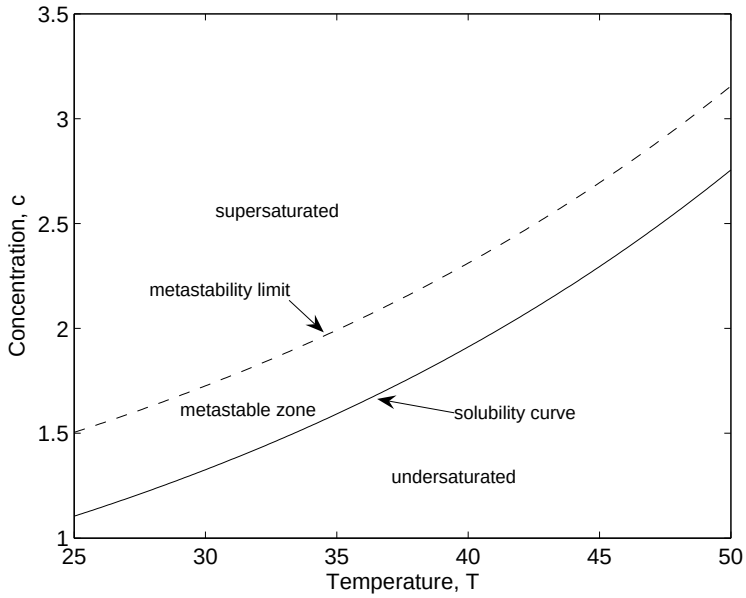


Figure 2.1: Solubility diagram.

2.1.2 Nucleation

The process of creating a new solid phase from a supersaturated homogeneous phase is called nucleation. Nucleation mechanisms are commonly lumped into one of two categories: primary and secondary nucleation, and can be further classified as shown in Figure 2.2 (adopted from [128]).

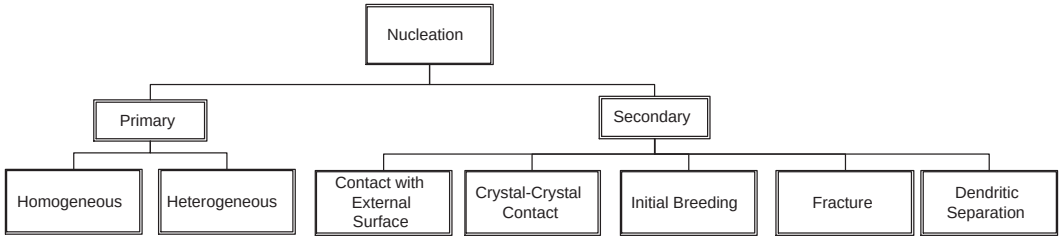


Figure 2.2: Nucleation mechanism.

Mechanism of formation of crystal that is independent of the presence of other suspended crystals is classified as primary or spontaneous nucleation. Primary nu-

cleation is associated with high levels of supersaturation and is usually partitioned as homogeneous nucleation, which occurs in the pure bulk solution, and heterogeneous nucleation, which is induced by foreign surfaces such as impurities.

Classical thermodynamic free energy minimization is used to derive the rate of homogeneous nucleation [163]. This theory postulates the production of embryos from the combination of solute molecules in a series of bimolecular reactions. The free energy of the embryos achieves a maximum at a critical size particular to the chemical system. Once an embryo exceeds this critical size, the free energy decreases with further growth, leading to spontaneous nucleation. The very high supersaturation required to overcome interfacial tension between embryo and solution makes homogeneous nucleation an unlikely mechanism for crystal formation for most chemical systems under industrial conditions [27, 163].

Foreign surfaces and particles promote nucleation as a result of an ordering process caused by interactions across the interface [163]. The result of this catalysis is that primary heterogeneous nucleation can occur at supersaturation levels significantly lower than required for homogeneous nucleation and is, therefore, the dominant mechanism of primary nucleation when impurities are present. Nevertheless, the supersaturation levels of heterogeneous nucleation are often still too high for good crystal growth and production of crystals of desirable morphology [27]. Also, classical theory suggests that primary heterogeneous nucleation is characterized by a process that is either starved for nuclei or overwhelmed by a burst of new crystals [27], making Crystal Size Distribution (CSD) control difficult.

Secondary nucleation describes the nucleation that takes place due to the presence of other solute crystals. It always accompanies primary nucleation and is the dominant mechanism in most industrial crystallizations [8, 51, 124]. Secondary nucleation is more easily controlled than primary nucleation and occurs at supersaturation levels conducive to good crystal quality. There are a variety of proposed mechanisms whereby the crystals promote formation of new crystals. For example, Botsaris [8] postulated two questions: what are the sources of the potential nuclei, and how are the potential nuclei extracted from the source and displaced into the bulk solution to initiate new crystals. Proposed sources of potential nuclei include protrusions from growing crystal surfaces, solute clusters on or near the crystal surface, and embryos in the supersaturated solution. Several mechanisms for the conversion of potential nuclei into nuclei have been proposed which include spontaneous removal of dendrites due to free energy driving force, fluid shear, and contact nucleation resulting from the contact of crystals with an external surface or other crystals [8, 27, 73, 125, 152].

The variety and complexity of the mechanisms of secondary nucleation have forced researchers to use the simplified modeling of nucleation by assuming an empirical functional form. The following expression is commonly used to describe secondary nucleation:

$$B^{\circ} = k_b \exp(-E_b/T) S^b \mu_k^j \,, \tag{2.1}$$

where k_b, b, E_b, and j are considered to be empirical constants, μ_k is the k^{th} moment of the CSD, and S is the relative supersaturation defined as $S = \frac{C-C_{\text{sat}}}{C_{\text{sat}}}$. The temperature effect on B° is complicated and some kinetic studies show that nucleation can actually decrease with increasing temperature, which corresponds to a negative activation energy [46]. A common hypothesis for this observation is that higher temperature leads to increased growth rates, involving greater efficiency of molecular diffusion on and integration into crystal surfaces; thus, fewer potential nuclei are available for secondary nucleation [60]. Nevertheless, evidence suggests that b is independent of temperature [46] and Arrhenius-type expression is probably adequate for characterizing temperature dependence.

2.1.3 Growth

As soon as stable nuclei (i.e. particles larger than the critical size) have been formed in a supersaturated or supercooled system, they begin to grow into crystals of visible size. At the microscopic level, solute molecules moving from the bulk solution adsorb on the crystal solid surface and are incorporated into the crystal lattice. A well-defined, smooth crystal face is planar and new solute molecules must migrate across the surface to find energetically favorable incorporation sites [128]. Three such potential sites are shown in Figure 2.3 (adopted from [128]). Site A is thermodynamically unfavorable compared with B, a step site or C, a kink site. Surface adsorption and diffusion determine whether a solute molecule is incorporated into the crystal or returns to the bulk phase [128].

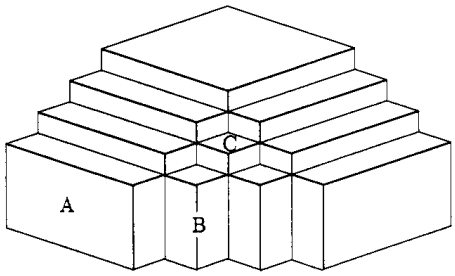


Figure 2.3: Crystal incorporation sites: flat faces (A), step sites (B), and kink sites (C).

Many attempts have been made to explain the mechanism and rate of crystal growth, which can be classified into three categories, namely, 'surface energy', 'diffusion' and 'adsorption-layer' theories [112].

The surface energy theories are based on the postulation of Gibbs (1878) and Curie (1885) that the shape a growing crystal assumes is that which has a minimum surface energy. This approach, although not completely abandoned, has largely fallen into disuse. The diffusion theories originated by Noyes and Whitney (1897) and Nernst (1904) presume that matter is deposited continuously on crystal face at a rate proportional to the difference in concentration between the point of deposition and the bulk of the solution. In 1922, Volmer suggested that crystal growth was a discontinuous process, taking place by adsorption, layer by layer, on the crystal surfaces [112].

For engineering purposes, the semiempirical power law has become the standard representation of the growth rate, which assumes the following form,

$$G = k_g \exp(-E_g/RT) L^{g_1} S^{g_2} . \tag{2.2}$$

where k_g, g_1, g_2, and E_g, are empirical constants.

Note that the temperature dependence of G has been incorporated using an Arrhenius-type expression. Also notice that the growth rate is dependent on the crystal size (L) by power law. If $g_1 = 0$, G is size-independent and this assumption is usually referred to as McCabe's ΔL law. However, there are several examples of systems that violate this assumption [17, 20]. Size-dependent growth rate is usually attributed to either bulk diffusion effects or the Gibbs-Thomson effect, which suggests an increasing growth rate with increasing size because of an inverse relationship between solubility and size [128]. Garside et al. [45] presents a theory of size-dependent surface integration kinetics.

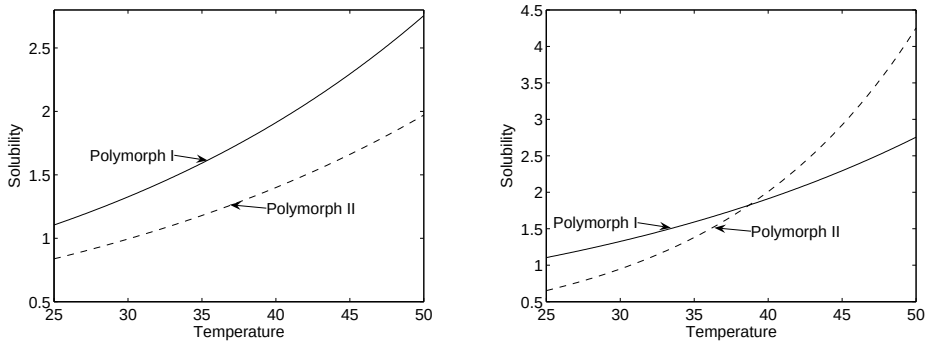
Modeling of growth rate is further complicated by a phenomenon known as growth rate dispersion. It describes the situation in which not all of the crystals grow at identical or constant rates although the crystallizer conditions remain constant. A more detailed discussion regarding growth rate dispersion can be found in [48, 71, 125, 126, 164, 176].

2.1.4 Polymorphism

The ability of a material to crystallize into more than one crystal form is known as polymorphism. Polymorphs of a given compound can have widely different properties such as dissolution rate, bioavailability, melting point, hardness and electrical properties. As a result, polymorphism is important in the pharmaceutical industry where safety and reliability are of paramount importance. The unexpected appear-

ance of a second polymorphic form of an active pharmaceutical ingredient used for the treatment of HIV, with substantially different dissolution and absorption characteristics, highlights the importance of polymorphism in the pharmaceutical industry [6]. Realizing the importance of polymorphism, the U.S. Food and Drug Administration (FDA) has tightened regulations for new drug applications to demonstrate control over the manufacturing process [16, 142].

In the chemical and pharmaceutical industry, the demand for high yields and high production rates has forced chemists and engineers to operate processes far from equilibrium, such that it exacerbates the tendency to form polymorphic structures. Hence, it is important to investigate the stability of each polymorph at a given temperature, pressure, and composition.



(a) Monotropic (b) Enantiotropic

Figure 2.4: Solubility curves in polymorphic systems.

The relative stability of each polymorph is reflected by its relative solubility, with the more stable polymorph having the lower solubility. In polymorphic systems, phase diagrams (i.e. solubility curves) are found to fall into one of the two

categories shown in Figure 2.4: Monotropic in which the relative stability of the polymorphs are independent of temperature, and Enantiotropic in which the relative stability of the polymorphs are temperature dependent [29].

The kinetic processes involved in phase transitions between polymorphs depend largely on the extent of structural changes involved [29]. Two possible transformations are

(1) Solvent-mediated (reconstructive) transformation, in which the metastable phase dissolves while the stable phase renucleates and grows from solution.

(2) Solid state (displacive) transformation, in which nucleation and growth of the new phase take place in crystals of the unstable phase. Such transitions are often reversible when the temperature is raised and lowered through the transition temperature

Solvent-mediated transformation often involves a lower activation energy than a solid state transformation. Consequently, the former is favored to occur well below melting point and therefore its kinetic mechanism is widely used in pharmaceutical industries.

2.2 Recent development on the modelling, simulation, and control of crystallization

2.2.1 Modelling

Together with the crystal morphology, the crystal size distribution (CSD) produced within crystallizer is of crucial importance in determining the ease and efficiency of subsequent solid-liquid separation steps, the suitability of crystals for further processing, their caking and storage characteristics, and the eventual customer acceptance of the product [29]. The most common way of tracking the CSD is by making use of population balance equations [125], which describe the material balance that accounts for the distribution of different size crystals in the crystallizer.

For simplification, most batch crystallization studies in the literature only consider nucleation and growth kinetics (i.e. no agglomeration and dendritic growth) and ignore shape changes. For non-polymorphic systems, the simplified population balance equation for one dimensional growth in a well-mixed batch process is given below:

$$\frac{\partial f}{\partial t} + \frac{\partial}{\partial L}\left(G(L, S, T; \theta_{\mathrm{g}}) f\right) = B(f, S, T; \theta_{\mathrm{b}}) \delta(L)\,, \tag{2.3}$$

whereas for polymorphic systems:

$$\frac{\partial f_i}{\partial t} + \frac{\partial}{\partial L}\left(G_i(L_i, S_i, T; \theta_{\mathrm{gi}}) f_i\right) + \frac{\partial}{\partial L}\left(D_i(L, S_i, T; \theta_{\mathrm{di}}) f_i\right) =$$

$$B_i(f_i, S_i, T; \theta_{\mathrm{bi}})\delta(L)\,, \tag{2.4}$$

where T is the temperature, S is the supersaturation, f_i (f) is the number density of polymorph i (crystals), L is the crystal's characteristic size, G_i (G) is the growth rate of polymorph i (crystals), D_i is the dissolution rate of the polymorph i, B_i (B) is the nucleation rate of polymorph i (crystals), δ is the Dirac delta function, θ_{gi} (θ_{g}), θ_{di}, and θ_{bi} (θ_{b}) are vectors of growth, dissolution and nucleation kinetic parameters of polymorph i (crystals), respectively.

The modelling of one dimensional and multidimensional non-polymorphic crystallization based on population balance equations has been discussed frequently in the literatures [54, 66, 67, 105, 109, 128, 158, 161]. Kinetics of polymorphic transformation process have been estimated by various procedures [21, 136, 137, 171], and weighted least squares method is commonly applied to estimate the model parameters [18, 40, 115, 139]. Although the papers by Ono et al. [115] and Scholl et al. [139] are major contributions to the field of polymorphic crystallization, the effect of temperature variation was neglected and the parameter uncertainties were not reported in these two papers. This motivates this study to develop a more comprehensive model which includes the effect of temperature variation and the marginal distributions of the kinetic parameters.

While weighted least squares methods are adequate for many problems, Bayesian inference (which will be employed in this study) is able to include prior knowledge in the statistical analysis which can produce models with higher predictive capa-

bility. In addition, the resulting posterior distribution for the estimated parameters can be used to accurately quantify the accuracy of model predictions and can be incorporated into robust control strategies for crystallization process [114]. Although Bayesian inference is not within the standard toolkit for chemical engineers, there have been many applications to chemical engineering problems over the years including the estimation of parameters in chemical reaction models [9], heat transfer in packed beds [38], microbial systems [7, 28, 117], and microelectronics processes [53].

2.2.2 Simulations

Numerical simulations of the resulting model enable the investigation of the effects of various operating conditions and can be used for optimal design and control [64, 130, 139]. Solving population balance equations is particularly challenging when the PDEs are hyperbolic with sharp gradients or discontinuities in the distribution [148]. Standard first-order methods require a very small grid size in order to reduce the numerical diffusion (i.e., smearing), whereas standard higher order methods introduce numerical dispersion (i.e., spurious oscillations), which usually results in a crystal size distribution with negative values. Therefore, efficient and sufficiently accurate computational methods for simulating the population balance equations are required to ensure the behaviour of the numerical solution is determined by the assumed physical principles and not by the chosen numerical method.

There have been many papers on the numerical solution of population balance

models (PBMs). The method of moments approximates the distribution by its moments [69], which under certain conditions, converts the hyperbolic PDEs into a small number of ordinary differential equations (ODEs) that describe characteristics of the distribution. The method of moments does not apply to PBEs which do not satisfy moment closure conditions. The method of weighted residuals approximates the size distribution by a linear combination of basis functions [147], which results in a system of ODEs. For most practical crystallizations, a large number of basis functions is needed to approximate the distribution, which results in high computational cost. The Monte Carlo method tracks individual particles, each of which exhibits stochastical behaviour according to a probabilistic model [15, 58, 123]. This approach is too computationally expensive for most industrial crystallizations. Another problem-specific numerical method for solving population balance equations is the method of characteristics [83, 121]. This method solves each population balance equation by finding curves in the characteristic size-time plane that reduce the equation to an ODE. While the method is highly efficient when the kinetics are simple, the approach does not generalize to more complex kinetics. Most publications on numerical methods for solving PBEs involve various types of discretizations and go by a variety of names including “method of classes” and “discretized population balance equations” [65, 82, 100, 101, 118]. In recent years there have been several efforts to reduce the numerical diffusion and numerical dispersion for distributions which contain sharp gradients or discontinuities, which is common in batch crystallizations. High resolution finite volume methods (FVMs) popular in

astrophysics and gas dynamics [56, 92, 93, 116, 153] were extended to the application of multidimensional population balance equations [52, 104, 119, 120, 166]. A typical implementation applies a first-order method near discontinuities or sharp gradients and a second-order method everywhere else, which results in less numerical dispersion than the second-order method and less numerical diffusion than the first-order method [52].

In recent years, new high order finite difference methods have been developed in the field of computational physics that are designed to robustly treat discontinuities by upwinding in the vicinity of a discontinuity while maintaining high order accuracy in smooth regions. Among those methods are the essentially non-oscillatory (ENO) finite difference methods [57]. A generalization and practical improvement of these very successful schemes is the weighted ENO (WENO) methods. In this study, various WENO methods [61, 72, 99, 141] are considered for solving population balance models. The performance of these WENO methods are compared to the high resolution (HR) finite volume method and a second-order finite difference (FD2) method, for the polymorphic crystallization model developed in this thesis.

2.2.3 Control

The vast majority of papers on non-polymorphic crystallization have considered the optimal control of only one or two characteristics of the crystal size distribution, such as weight mean size. The most widely studied approach is to determine a temperature profile (nominal T-control) that optimizes an objective function based on

an offline nominal model [66, 86, 128, 167, 174]. Then, this temperature profile is used as the recipe during the crystallization process. One approach to implement this control strategy is to parameterize the temperature-time trajectory into piecewise linear functions with temperature values at some points of time as the decision variables. Then, optimization is carried out to minimize/maximize a specified objective function. Although T-control is simple to implement, it has become well-known in recent years that T-control can be very sensitive to variations in the kinetic parameters resulting from plant-model mismatch [13, 133].

This motivated the development of robust T-control [36, 102, 149]. This approach is similar to the nominal T-control, with the objective function explicitly includes the impact of uncertainties while determining the optimal temperature-time trajectory to be followed during batch operation. One approach to include the impact of uncertainties is through the worst-case analysis, where worst-case parameter vector and an initial estimate of performance degradation based on a Taylor series expansion that describes the local behavior about the nominal trajectory is computed first. Then, a nonlinear dynamic simulation is used to compute the improved estimates. The difference between the initial and final estimates provides an indication as to the accuracy of the Taylor series expansion in capturing the process dynamics in the vicinity of the control trajectory. These quantitative estimates can be used to decide whether more laboratory experiments are needed to produce parameter estimates of higher accuracy, or to define performance objectives for lower-level control loops that implement the optimal control trajectory [103].

With advances in sensor technologies, another control strategy developed to provide improved robustness to model uncertainty is C-control, which follows an optimal or nearly optimal concentration-temperature trajectory [41, 42, 50, 97, 133, 175]. There are two common approaches in implementing the C-control, namely, the first-principles approach and the direct-design approach. In the former approach, a model constructed from material and energy balances are used to optimize an objective function, in which the decision variables comprises the parameters of the concentration-temperature trajectory parameterization. Then, the resulting concentration-temperature trajectory is used as the setpoint for the lower-level control loop during the crystallization process. On the other hand, the latter approach does not require any kinetics model and determines a suboptimal concentration-temperature trajectory within the metastable zone as the setpoint to be followed during the crystallization process. In many experimental and simulation studies of non-polymorphic batch crystallizations, the C-control strategy has resulted in low sensitivity of the product quality to most practical disturbances and variations in kinetic parameters. Recently, the C-control strategy has been applied to polymorphic crystallizations to produce large crystals of any selected polymorph [80] and to ensure maximum productivity in polymorphic transformation process [64]. In this thesis, both T-control and C-control strategies will be applied to the polymorphic crystallization process model developed in this study. In addition, their resulting control performance will serve as benchmark for the other control techniques developed in this thesis.

Despite the high impact of model predictive control (MPC) [19, 34, 44, 62, 110, 122, 127] in academic research and industrial practice, its application to solution crystallization processes has been rather limited [35, 79, 113, 131, 155]. One contribution considered the effects of uncertainties on the closed-loop performance of nonlinear model predictive control (NMPC) applied to crystallization processes [113]. As in many other papers, the method of moments was utilized to simplify the population balance equations which are partial different equations (PDEs) to a set of ordinary differential equations (ODEs) in terms of the moments. The NMPC optimization problem was solved using nonlinear programming and the states were estimated using an extended Kalman filter (EKF). To the author's knowledge, there is no published result on the implementation of NMPC to polymorphic crystallization, which is more challenging for a number of reasons. First, the phase equilibria and crystallization kinetics are more complicated. Second, the method of moments heavily used in past control algorithms for crystallization processes does not apply during a polymorphic transformation, so that the full PDEs need to be solved. As a consequence, the computation time required increases considerably which prohibits the straightforward application of nonlinear programming. To alleviate this shortcoming, a practical NMPC control strategy for the polymorphic crystallization will be developed in this thesis.

Recognizing that the system under study is a batch process, which is repetitive in nature, it is possible to utilize the information from the previous batches to improve the control performance from batch-to-batch (batch-to-batch control). The key idea

of batch-to-batch control is to improve the transient response performance of an unknown/uncertain system that operates repetitively over a fixed time interval by using the previous actual operation data to compensate for uncertainty [2]. Though batch-to-batch control strategy has been widely studied in major chemical processes such as polymerisation process [26, 37, 169], rapid thermal processing [88], and so on, the application to crystallization processes has not been found. This serves as a motivation to investigate the implementation of batch-to-batch control to the polymorphic crystallization process under study.

Chapter 3

Modelling the Crystallization of L-glutamic Acid Polymorphs

3.1 Introduction

In this chapter, a kinetic model of L-glutamic acid polymorphic crystallization is developed from batch experiments with in-situ measurements including attenuated total reflection Fourier transform infrared (ATR-FTIR) spectroscopy to infer the solute concentration and focused beam reflectance measurement (FBRM) which provides crystal size information. Kinetics of polymorphic transformation have been estimated by various procedures [21, 136, 137, 171]. A commonly used method to estimate model parameters in nonlinear process models is weighted least squares [3, 4, 106], which has been applied to polymorphic crystallization [18, 40, 115, 139]. While weighted least squares methods are adequate for many

problems, Bayesian inference is able to include prior knowledge in the statistical analysis which can produce models with higher predictive capability. Although Bayesian inference is not within the standard toolkit of chemical engineers, there have been many applications to chemical engineering problems over the years including the estimation of parameters in chemical reaction models [9], heat transfer in packed beds [38], microbial systems [7, 28, 117], and microelectronics processes [53].

Quantifying uncertainties in the parameter estimates is required for assessing the accuracy of model predictions [109, 114]. When weighted least squares methods are used for parameter estimation, the widely used approaches to quantify uncertainties in parameter estimates are the linearized statistics and likelihood ratio approaches [5]. In the linearized statistics approach, the model is linearized around the optimal parameter estimates and the parameter uncertainty is represented by a χ^2 distribution. This model linearization can result in highly inaccurate uncertainty estimates for highly nonlinear models [5], and this approach ignores physical constraints on the model parameters. The likelihood ratio approach, which is the nonlinear analogue to the well-known F statistic, takes nonlinearity into account but approximates the distribution [5], and ignores constraints on the model parameters. This study applies a Bayesian inference approach that not only avoids making these approximations but also includes prior information during the estimation of parameter uncertainties.

In this study, the parameters in a kinetic model for L-glutamic acid polymor-

phic crystallization process are determined by Bayesian estimation. The probability distribution of process model parameters is defined through the Bayesian posterior density, from which all parameter estimates of interest (e.g., means, modes, and credible intervals) are calculated. However, the conventional approach to calculate the above estimates often involves complicated integrals of the Bayesian posterior density which are analytically intractable. To overcome this drawback, Markov Chain Monte Carlo (MCMC) integration [47, 98, 157] was applied to compute these integrals in an efficient manner. MCMC does not require approximation of the posterior distribution by a Gaussian distribution [23, 28, 85]. This posterior distribution for the estimated parameters can be used to accurately quantify the accuracy of model predictions and can be incorporated into robust control strategies for crystallization process [114].

This chapter is organized as follows. The next section describes the experimental procedure to obtain measurement data for parameter estimation. A short review of Bayesian theory and MCMC integration is in Section 3.3. In Section 3.4, the L-glutamic acid crystallization model is described and the parameter estimation results discussed. This is followed by the conclusions.

3.2 Experimental methods

The crystallization instrument setup used was similar to that described previously [41]. A Dipper-210 ATR immersion probe (Axiom Analytical) with ZnSe as the

Table 3.1: L-glutamic acid aqueous solutions used for calibration.

Concentration $[\mathrm{g/g\,of\,water}]$	Temperature range $[^{\circ}\mathrm{C}]$
0.00837	35 to 21
0.01301	48 to 13
0.01800	57 to 32
0.02300	64 to 34
0.02800	64 to 45

internal reflectance element attached to a Nicolet Protege 460 FTIR spectrophotometer was used to obtain L-glutamic acid spectra in aqueous solution, with a spectral resolution of 4 cm^{-1}. The chord length distribution (CLD) for L-glutamic crystals in solution were measured using Lasentec FBRM connected to a Pentium III running version 6.0b12 of the FBRM Control Interface software.

3.2.1 Calibration for solution concentration

Different solution concentrations of L-glutamic acid (99%, Sigma Aldrich) and degassed deionized water were placed in a 500-ml jacketed round-bottom flask and heated until complete dissolution. The solution was then cooled at 0.5 $^{\circ}\mathrm{C}/\mathrm{min}$ while the IR spectra were being collected, with continuous stirring in the flask using an overhead mixer at 250 rpm. Table 3.1 lists the five different solution concentrations used to build the calibration model.

The IR spectra of aqueous L-glutamic acid in the range 1100-1450 cm^{-1} and the temperature were used to construct the calibration model based on various chemometrics methods such as principal component regression (PCR) and partial least squares regression (PLS) [159]. The calculations were carried out using in-house

MATLAB 5.3 (The Mathworks, Inc) code except for PLS, which was from the PLS Toolbox 2.0. The mean width of the prediction interval was used as the criterion to select the most accurate calibration model. The noise level was selected based on the compatibility of the prediction intervals with the accuracy of the solubility data. The chemometrics method forward selection PCR 2 (FPCR 2) [168] was selected because it gave the smallest prediction interval; using a noise level of 0.001, the prediction interval (0.73 g/kg) was compatible with the accuracy of this model with respect to solubility data reported in the literature [115].

3.2.2 Solubility determination and feedback concentration control experiments

The commercially available L-glutamic acid crystals were verified to be pure β-form using powder X-ray diffraction (XRD) and were used for the determination of the β-form solubility curve. Pure α-form crystals obtained using a rapid cooling method outlined previously [115] were used to determine the α-form solubility curve in similar fashion as the β-form in a separate experiment. For each polymorph, the IR spectra of L-glutamic acid slurries (saturated, and with excess crystals) were collected at different temperatures ranging from $25^{\circ}\mathrm{C}$ to $60^{\circ}\mathrm{C}$. The slurry was equilibrated for 45 minutes to 1 hour at a specified temperature before recording the IR spectra. The solution concentration was then calculated using the aforementioned calibration model. The resulting solubility measurements for L-glutamic acid

Table 3.2: Solubility data for L-glutamic acid polymorphs.

Temperature $[^\circ\mathrm{C}]$	Solubility of α-form $[\mathrm{g/kg}]$	Solubility of β-form $[\mathrm{g/kg}]$
25	10.5971	8.5434
30	13.1599	9.7362
35	15.8004	12.4257
40	19.1689	13.7163
45	23.3185	17.0729
50	27.0364	19.8722
55	31.7768	23.3904
60	36.8028	27.7567

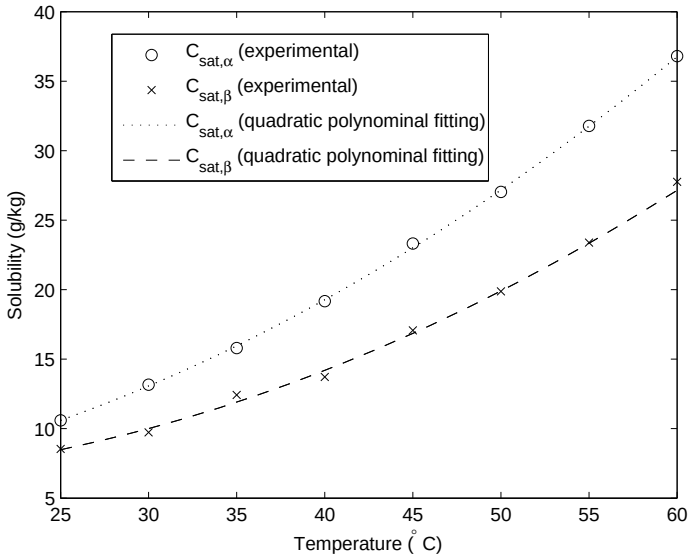


Figure 3.1: Solubility curves of L-glutamic acid polymorphs.

polymorphs are tabulated in Table 3.2 and Figure 3.1 compares the measurements to their quadratic polynomial fitting.

In the seeded batch crystallization experiments, appropriate amounts of L-glutamic acid (99%, Sigma Aldrich) in 400 g of water was heated to about $5^\circ\mathrm{C}$ above the β-form saturation temperature in a 500-mL jacketed round-bottom flask with an overhead mixer at 250 rpm, to create an undersaturated solution. The crystallizer was

then cooled and seed crystals (either pure α- or β-form) were added when the solution was supersaturated with respect to the seeded form. Different supersaturation setpoint profiles were followed during crystallization based on in-situ solution concentration measurement as described previously [41]. The control algorithm was started shortly after seeding.

3.3 Review of Bayesian inference

3.3.1 Bayesian posterior

Bayesian inference is the process of fitting a probability model to a set of data and summarizing the results by a probability distribution on the parameters of the model and on unobserved quantities such as predictions for new observations [47]. The fundamental difference between Bayesian and traditional statistical methods is the interpretation of probability. Classical methods, also known as the frequentist methods, perceive probability as the long-run relative frequency of occurrence determined by the repetition of an event. A Bayesian method perceives probability as a quantitative description of the degree of belief in a given proposition [14, 28]. With this interpretation of probability, the Bayesian method allows a practitioner to account for prior information in a statistical analysis.

Furthermore, Bayesian inference facilitates a common-sense interpretation of statistical conclusions. For instance, a Bayesian credible interval for an unknown quantity of interest can be directly regarded as having a high probability of con-

taining the unknown quantity, in contrast to a frequentist confidence interval, which may strictly be interpreted only in relation to a sequence of similar inferences that might be made in repeated practice. A brief introduction to Bayesian inference is given below. Interested readers are referred to references [14, 22, 47] for a thorough discussion.

The main substance of Bayesian inference is Bayes' rule:

$$\Pr(\theta|\mathbf{y}) = \frac{\Pr(\mathbf{y}|\theta)\Pr(\theta)}{\Pr(\mathbf{y})}\,, \tag{3.1}$$

where θ is a vector of unknown parameters of interest and $\mathbf{y}$ represents the collected data which is used to infer θ. These data usually consist of observed state variables (e.g., concentration) at different time points. $\Pr(\theta)$ is the prior distribution of θ, $\Pr(\mathbf{y}|\theta)$ is referred as the sampling distribution (or data distribution) for fixed parameters θ. When the data $\mathbf{y}$ are known and the parameters θ are unknown (i.e., as parameter estimation), the term $\Pr(\mathbf{y}|\theta)$ is referred as the likelihood function and denoted as $\mathrm{L}(\theta|\mathbf{y})$. $\Pr(\theta|\mathbf{y})$ is referred as the Bayesian posterior distribution of θ, and $\Pr(\mathbf{y}) = \int \Pr(\mathbf{y}|\theta)\Pr(\theta)d\theta$ acts as a normalizing constant to ensure that the Bayesian posterior integrates to unity. This constant is also called marginal likelihood or Bayes factor. For the inference of θ, the Bayes factor can be omitted since it does not affect the the resulting posterior distribution of θ, which yields the

unnormalized posterior distribution:

$$\Pr(\theta|\mathbf{y}) \propto \mathrm{L}(\theta|\mathbf{y})\Pr(\theta)\,. \tag{3.2}$$

In this study, it is assumed that the model structure is correct, and the measurement noise is distributed normally with zero mean and unknown variance. Then the likelihood is of the form

$$\begin{aligned}
\mathrm{L}(\theta|\mathbf{y}) &= \mathrm{L}(\theta_{\mathrm{sys}}, \sigma|\mathbf{y}) \\
&= \prod_{j=1}^{N_m}\prod_{k=1}^{N_{d_j}} \Pr(y_{jk}|\theta_{\mathrm{sys}}, \sigma) \\
&= \prod_{j=1}^{N_m}\prod_{k=1}^{N_{d_j}} \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left(-\frac{\left(y_{jk} - \hat{y}_{jk}(\theta_{\mathrm{sys}})\right)^2}{2\sigma_j^2}\right) \\
&= \frac{1}{\prod_{j=1}^{N_m}\left(\sqrt{2\pi}\sigma_j\right)^{N_{d_j}}} \exp\left(-\sum_{j=1}^{N_m}\sum_{k=1}^{N_{d_j}} \frac{\left(y_{jk} - \hat{y}_{jk}(\theta_{\mathrm{sys}})\right)^2}{2\sigma_j^2}\right)\,,
\end{aligned} \tag{3.3}$$

where $\theta = [\theta_{\mathrm{sys}}, \sigma]^T$ is the vector of parameters of interest which consist of system/model (θ_{sys}) and noise (σ) parameters, y_{jk} and $\hat{y}_{jk}$ are the measurement and predicted value of j^{th} variable at sampling instance k, respectively, N_m is the number of measured variables, N_{d_j} is the number of time samples of j^{th} variable, and σ_j is the standard deviation of the measurement noise in the j^{th} variable.

The prior distribution $\Pr(\theta)$ can be informative or non-informative, depending on the prior knowledge of θ. The most commonly used non-informative prior is $\Pr(\theta) \propto 1$. However, this is an improper prior distribution, since its integral is in-

finity, and may lead to an improper posterior distribution. The use of an informative prior distribution is preferred, for example, a prior distribution which specifies the minimum and maximum possible values of θ is

$$\Pr(\theta) \propto \begin{cases} 1 & \text{if } \theta_{\min} \leq \theta \leq \theta_{\max} \\ 0 & \text{otherwise,} \end{cases} \tag{3.4}$$

which means that all values of θ between $\theta_{\min}$ and $\theta_{\max}$ have equal probability. In cases where the prior distribution is available from past parameter estimation studies, the distribution is not uniform [53]. A detailed discussion regarding informative and non-informative priors can be found in the literature [10, 22, 47].

The product of the likelihood and prior distribution defines the Bayesian posterior, which is the joint probability distribution for all parameters after data have been observed. Once the Bayesian posterior is defined, it is desirable to determine the mean, mode, and credible intervals associated with each of the parameters. Markov chain simulation, also called Markov Chain Monte Carlo (MCMC), is employed for that purpose in this study.

3.3.2 Markov chain simulation

Markov chain simulation draws values of θ from approximate distributions and then corrects these values to better approximate the target distribution. In this case, the target distribution is the Bayesian posterior. The samples are drawn sequentially,

with the distribution of the sampled values depending on the last value drawn. The Markov chain is a sequence of random variables θ^0, θ^1, . . ., for which, for any s, the distribution of θ^{s+1} given all previous θ's depends only on the most recent value, θ^s. The key to the method's success, however, is not the Markov property but rather that the approximate distributions are improved at each step in the simulation, in the sense of converging to the target distribution.*

In the application of Markov chain simulation, several parallel chains are drawn. Parameters from each chain c, $\theta^{c,s}$, $s = 1, 2, 3, \ldots,$ are produced by starting at some point $\theta^{c,0}$ and then, for each step s, drawing $\theta^{c,s+1}$ from a jumping distribution, $T_s(\theta^{c,s+1}|\theta^{c,s})$ that depends on the previous draw, $\theta^{c,s}$. The jumping probability distributions must be constructed so that the Markov chain converges to the target posterior distribution.

The Metropolis algorithm [108] is a simple algorithm to construct a Markov chain which converges to the posterior distribution. The algorithm is an adaptation of a random walk that uses an acceptance/rejection rule to converge to the specified target distribution. In the Metropolis algorithm, the widely used approach to create the next step of the chain c, θ^{c,s_p}, is to perturb the current step of the chain $\theta^{c,s}$ by adding some amount of noise ($\theta^{c,s_p} = \theta^{c,s} + \epsilon$), where ϵ is distributed normally with zero mean and covariance matrix Σ. However, specifying the covariance matrix can be challenging. This covariance matrix needs to be chosen in such a way so as to balance progress in each step and a reasonable acceptance rate. A poorly chosen co-

*For further information on Markov chains, readers are referred to other literature [47, 98].

variance matrix may cause slow convergence. Traditionally, the covariance matrix is estimated from a trial run and much recent research is devoted to ways of doing that efficiently and/or adaptively [55]. If parameters θ are highly correlated, special precautions must be taken to avoid singularity of the estimated covariance matrix.

Recently, there has been a development in combining evolutionary algorithms with MCMC. [11, 87, 95, 96]. Among others, the combination of differential evolution (DE) with MCMC is particularly interesting. Behind DE is an evolutionary algorithm for numerical optimization; its combination with MCMC (shortened as DE-MC [11]) solves an important problem in MCMC, namely that of choosing an appropriate scale and orientation for the jumping distribution (i.e., related to the covariance matrix Σ in the Metropolis algorithm). In DE-MC, the jumps are simply a fixed multiple of two random parameter vectors that are currently in the population, and the selection process of DE-MC works via the usual Metropolis ratio which defines the probability with which a proposal is accepted. Motivated by its efficiency and effectiveness, DE-MC is utilized to construct the Markov chains of θ in this study.

Constructing the Markov chains is one step. Next is to monitor the convergence of the chains in order to decide how many samples need to be collected or when to stop the MCMC simulation. Too few samples will result in an inaccurate distribution of the parameters θ. Here, potential scale reduction factors ($\hat{R}_i$) were adopted to monitor the convergence of the Markov chains [47], which estimate the potential improvement in the Markov chain estimation of the respective i^{th} parameter θ_i if

the Markov chain simulation were continued. This potential scale reduction factor is calculated from the following equations:

$$\hat{R}_i = \sqrt{\frac{\hat{\mathrm{var}}^+(\theta|\mathbf{y})_\mathrm{i}}{W_i}}\,, \tag{3.5}$$

$$\hat{\mathrm{var}}^+(\theta|\mathbf{y})_\mathrm{i} = \frac{n-1}{n}W_i + \frac{1}{n}B_i\,, \tag{3.6}$$

$$B_i = \frac{n}{m-1}\sum_{c=1}^{m}\left(\bar{\theta}_i^c - \bar{\theta}_i\right)^2\,, \tag{3.7}$$

$$W_i = \frac{1}{m}\sum_{c=1}^{m}\left(d_i^c\right)^2\,, \tag{3.8}$$

$$\bar{\theta}_i^c = \frac{1}{n}\sum_{s=1}^{n}\theta_i^{c,s}\,, \tag{3.9}$$

$$\bar{\theta}_i = \frac{1}{m}\sum_{c=1}^{m}\bar{\theta}_i^c\,, \tag{3.10}$$

$$\left(d_i^c\right)^2 = \frac{1}{n-1}\sum_{s=1}^{n}\left(\theta_i^{c,s} - \bar{\theta}_i^c\right)^2\,, \tag{3.11}$$

where $\theta_i^{c,s}$ is the simulation draws of parameter i from step chain c at step s, B_i and W_i are the between- and within-sequence variances of parameter i, respectively, m is the number of parallel chains, with each chain of length n. The potential scale reduction factor decreases asymptotically to 1 as $n \to \infty$. Once $\hat{R}_i$ is near 1[†] for all i, it is safe to stop the simulation.

To summarize, the following is the procedure for constructing Markov chains using DE-MC with the potential scale reduction factor as the stopping criterion:

(1) Draw starting parameters for all chains, $\theta^{c,0}$ ($c = 1, \ldots, m$), from a starting distribution or choose starting parameters from dispersed values around a

[†]According to Gelman et al. [47], a value below 1.1 is acceptable.

crude approximation of the parameters.

(2) At each step, create a proposed value θ^{c,s_p} according to the jumping rule

$$\theta^{c,s_p} = \theta^{c,s} + \gamma\left(\theta^{R_1,s} - \theta^{R_2,s}\right) + \mathbf{e}\,, \tag{3.12}$$

where $\mathbf{e}$ is drawn from a symmetric distribution with a small variance compared to that of the target, but with unbounded support (e.g., $\mathbf{e} \sim N(0,b)^{N_\theta}$ with b small, $b = 10^{-4}$ is utilized in this study), N_θ is the number of parameters in θ, $\theta^{R_1,s}$ and $\theta^{R_2,s}$ are randomly selected without replacement from all chains at step s, and γ is a scaling constant with typical values between 0.4 and 1. From the guidelines in the literature [11], the optimal choice of γ is $2.38/\sqrt{2N_\theta}$. This choice of γ is expected to give an acceptance probability of 0.44 for $N_\theta = 1$, 0.28 for $N_\theta = 5$, and 0.23 for large N_θ.

(3) Calculate the ratio of the posterior densities,

$$r = \frac{\Pr\left(\theta^{c,s_p}|\mathbf{y}\right)}{\Pr\left(\theta^{c,s}|\mathbf{y}\right)}\,. \tag{3.13}$$

and obtain $\theta^{c,s+1}$ from

$$\theta^{c,s+1} = \begin{cases} \theta^{c,s_p} & \text{with probability } \min\{r,1\} \\ \theta^{c,s} & \text{otherwise.} \end{cases} \tag{3.14}$$

(4) For each parameter i, calculate the potential scale reduction factor $\hat{R}_i$ by Eqs. (3.5) to (3.11). If $\hat{R}_i \leq 1.1$ for all $i = 1, 2, \ldots, N_\theta$, stop the iteration and construct the matrix

$$\Theta = \begin{bmatrix} \theta_1^1 & \cdots & \theta_{N_\theta}^1 \\ \vdots & \ddots & \vdots \\ \theta_1^{N_s} & \cdots & \theta_{N_\theta}^{N_s} \end{bmatrix}, \tag{3.15}$$

where Θ contains the approximated samples from the target distribution and N_s is the total number of values drawn from the second halves for all the chains.

Otherwise, if $\hat{R}_i > 1.1$ for any i, set $s = s + 1$ and go to Step 2.

3.3.3 Monte Carlo integration

In the previous sections, the Bayes posterior was defined and a method for drawing samples from it was described, from which a matrix Θ was generated. Here, the significance of this matrix is described through its use for calculating the desired properties of the Bayes posterior.

In order to calculate any properties of the Bayes posterior, it is necessary to evaluate integral

$$E[f(\theta)] = \int_{\theta_{\min}}^{\theta_{\max}} f(\theta) \Pr(\theta|\mathbf{y}) d\theta \,, \tag{3.16}$$

where $E[\cdot]$ is the expected value, $f(\theta)$ is a function for which the expected value is to be estimated. Conventionally, this integration can be performed analytically if the resulting function inside the integral operator is simple. However, the Bayesian posteriors most often have irregular forms such that analytical integrations become infeasible. In such situations, it is suitable to perform Monte Carlo integration [47, 98, 157] which utilizes the matrix Θ obtained in the previous section:

$$\begin{aligned} E[f(\theta)] &= \lim_{\mathrm{N_s}\to\infty} \frac{1}{N_s} \sum_{l=1}^{N_s} f(\theta^l) \\ &\approx \frac{1}{N_s} \sum_{l=1}^{N_s} f(\theta^l) \ \text{ for large } N_s \,, \end{aligned} \tag{3.17}$$

where $\theta^l = \left[\theta_1^l, \theta_2^l, \ldots, \theta_{N_\theta}^l\right]$ is a random sample drawn from the Bayesian posterior which is obtained from the l^{th} row of matrix Θ. For example, the mean of each parameter θ_i is obtained by setting $f(\theta^l) = \theta^l$ in Eq. (3.17).

It is also desirable to obtain the marginal mode and credible interval for each parameter. Conventionally, this is done by drawing samples from the marginal posterior for each parameter and analyzing their histograms, where the marginal posterior is calculated by integrating the Bayes posterior with respect to all parameters except the desired parameter as follows

$$\begin{aligned} \Pr(\theta_i|\mathbf{y}) = \int_{\theta_{1,\min}}^{\theta_{1,\max}} \cdots \int_{\theta_{j,\min}}^{\theta_{j,\max}} \cdots \int_{\theta_{N_\theta,\min}}^{\theta_{N_\theta,\max}} \Pr(\theta_1, \ldots, \theta_j, \ldots, \theta_{N_\theta}|\mathbf{y}) \\ d\theta_1 \cdots d\theta_j \cdots d\theta_{N_\theta} \,, \end{aligned} \tag{3.18}$$

where $j \neq i$ and $\Pr(\theta_i|\mathbf{y})$ is the marginal posterior of θ_i. By taking advantage of the MCMC approach, this integration is not required since the samples from the marginal posterior of θ_i are given by the i^{th} column of the matrix Θ. The marginal mode of θ_i was estimated by determining the highest peak in the histograms of the marginal posterior. Finally, the 95% credible interval of θ_i was estimated by determining the range of θ_i which have cumulative marginal distribution between 2.5% to 97.5%.

3.4 L-glutamic acid crystallization model

A kinetic model for the crystallization of metastable α-form and stable β-form crystals of L-glutamic acid is developed. This appears to be the first model for polymorphic crystallization that includes all of the kinetic processes, and also includes their dependence on the temperature. An earlier model for this system did not include the nucleation and growth kinetics of α-form crystals [115]. An improved model which includes those kinetics [139] only considered primary heterogeneous nucleation, which only applies when the crystallization is either starved of nuclei or overwhelmed by a burst of new crystals, and hence not applicable to industrial practice [27]. To develop a model amenable for industrial application, secondary nucleation is considered in this study.

3.4.1 Kinetic model

The mass balance on the crystals is described by a population balance equation [69]

$$\frac{\partial f_i}{\partial t} + \frac{\partial (G_i f_i)}{\partial L} = B_i \delta(L - L_0)\,, i = \alpha, \beta \tag{3.19}$$

where f_i is the crystal size distribution of the i-form crystals $[\#/\mathrm{m}^4]$ (i.e., α- or β-form crystals), B_i and G_i are the nucleation $[\#/\mathrm{m}^3\mathrm{s}]$ and growth rate $[\mathrm{m}/\mathrm{s}]$ of the i-form crystals, respectively, L and L_0 are the characteristic size of crystals $[\mathrm{m}]$ and nuclei $[\mathrm{m}]$, respectively, and $\delta(\cdot)$ is a Dirac delta function.

For parameter estimation, the method of moments[‡] was applied to Eq. (3.19) to give

$$\frac{d\mu_{i,0}}{dt} = B_i\,, \tag{3.20}$$

$$\frac{d\mu_{i,n}}{dt} = nG_i\mu_{i,n-1} + B_i L_0^n, \quad n = 1, 2, \ldots\,, \tag{3.21}$$

where the nth moment of the i-form crystals $[\#\,\mathrm{m}^{n-3}]$ is given by

$$\mu_{i,n} = \int_0^\infty L^n f_i dL\,. \tag{3.22}$$

[‡]The approach applies for the experimental conditions in this study in which data were collected during nucleation and growth. The full population balance equation (3.19) is used under conditions in which dissolution occurs.

The above equations are augmented by the solute mass balance:

$$\frac{dC}{dt} = -3\frac{10^3}{\rho_{\text{solv}}}\left(\rho_\alpha k_{v\alpha} G_\alpha \mu_{\alpha,2} + \rho_\beta k_{v\beta} G_\beta \mu_{\beta,2}\right) , \tag{3.23}$$

where C is the solute concentration $[\text{g/kg}]$, ρ_{solv} is the density of the solvent $[\text{kg/m}^3]$, ρ_i is the density of the i-form crystals $[\text{kg/m}^3]$, k_{vi} is the volumetric shape factor of the i-form crystals (dimensionless) as defined by $v_i = k_{vi}L^3$, where v_i is the volume of the i-form crystal $[\text{m}^3]$, and 10^3 is a constant $[\text{g/kg}]$ to ensure unit consistency. The kinetic expressions are

$$B_\alpha = k_{b\alpha}(S_\alpha - 1)\mu_{\alpha,3} \tag{3.24}$$

(α-form crystal nucleation rate) ,

$$G_\alpha = \begin{cases} k_{g\alpha}(S_\alpha - 1)^{g_\alpha} & \text{if } S_\alpha \geqslant 1 \\ k_{d\alpha}(S_\alpha - 1) & \text{otherwise} \end{cases} \tag{3.25}$$

(α-form crystal growth/dissolution rate) ,

$$B_\beta = k_{b\beta,1}\left(S_\beta - 1\right)\mu_{\alpha,3} + k_{b\beta,2}\left(S_\beta - 1\right)\mu_{\beta,3} \tag{3.26}$$

(β-form crystal nucleation rate) ,

$$G_\beta = k_{g\beta,1}\left(S_\beta - 1\right)^{g_\beta} \exp\left(-\frac{k_{g\beta,2}}{S_\beta - 1}\right) \tag{3.27}$$

(β-form crystal growth rate) ,

where $S_i = C/C_{\text{sat},i}$ and $C_{\text{sat},i} = a_{i,1}T^2 + a_{i,2}T + a_{i,3}$ are the supersaturation and the saturation concentration $[\text{g/kg}]$ of the i-form crystals, respectively, and T is the

solution temperature $[^{\circ}\mathrm{C}]$. The kinetic parameters $k_{b\alpha}$, $k_{g\alpha}$, and $k_{d\alpha}$ correspond to the nucleation $[\#/\mathrm{m}^3\mathrm{s}]$, growth $[\mathrm{m/s}]$, and dissolution $[\mathrm{m/s}]$ rates of α-form crystals, respectively, whereas $k_{b\beta,j}$ and $k_{g\beta,j}$ correspond to the $j\mathrm{th}$ nucleation $[\#/\mathrm{m}^3\mathrm{s}]$ and growth $[\mathrm{m/s}]$ for $j = 1$ and dimensionless for $j = 2$ rates of β-form crystals, respectively, and g_i is the growth exponent of the i-form crystals which may have a value between 1 (for diffusion-limited growth) and 2 (for surface integration-limited growth) [107]. The Arrhenius equation was used to account for the variability of crystal growth rate with temperature:

$$k_{g\alpha} = k_{g\alpha,0} \exp\left(-\frac{E_{g\alpha}}{8.314\,(T+273)}\right), \tag{3.28}$$

$$k_{g\beta,1} = k_{g\beta,0} \exp\left(-\frac{E_{g\beta}}{8.314\,(T+273)}\right), \tag{3.29}$$

where $k_{gi,0}$ and E_{gi} are the pre-exponential factor $[\mathrm{m/s}]$ and activation energy $[\mathrm{J/mol}]$ for the growth rate of i-form crystals, respectively. The values for densities, volumetric shape factors, and parameters for the saturation concentration are in Table 3.3.

Secondary nucleation is assumed for both α- and β-form crystals, since it is the dominant nucleation process in seeded crystallization. Primary nucleation is not included in the model since it is negligible compared to the secondary nucleation. The nucleation rate expression (3.25) and the second term in Eq. (3.27) were adapted from that reported in the literature for β crystals for L-glutamic acid [115]. We have introduced the first term in Eq. (3.27) to model the nucleation of β-form crystals

Table 3.3: Values for densities, volume shape factors, and saturation concentration parameters.

Parameters	Values
ρ_{solv}	990
ρ_α	1540
ρ_β	1540
$k_{v\alpha}$	0.480
$k_{v\beta}$	0.031
$a_{\alpha,1}$	8.437×10^{-3}
$a_{\alpha,2}$	0.03032
$a_{\alpha,3}$	4.564
$a_{\beta,1}$	7.644×10^{-3}
$a_{\beta,2}$	-0.1165
$a_{\beta,3}$	6.622

from the surface of α-form crystals. The growth rate expression for the α-form crystals includes both growth (positive supersaturation) and dissolution (undersaturation). Dissolution occurs during the polymorphic transformation of α- to β-form crystals, where α-form crystals dissolve and β-form crystals nucleate and grow. As reported in the literature [115, 139], the dissolution kinetics cannot be estimated accurately from polymorphic transformation experiments, as the growth rate of β-form crystals is limiting. Thus the simple form of dissolution rate with exponential factor of 1 was used with $k_{d\alpha}$ determined by a correlation equation based on mass transfer-limited dissolution, as reported in the literature [139]. The growth rate expressions for both α- and β-form crystals are also adopted from the literature [81], except that the exponential term for the α-form crystals is omitted in this study as it had a negligible effect on the model fitness to the data.

3.4.2 Parameter estimation

Before parameter estimation is carried out, the measured variables are discussed first. The various in-situ sensors that have become available for crystallization processes have removed or reduced sampling of the crystal slurry during crystallization and reduced the amount of pharmaceutical needed for each batch experiment. The two in-situ measurements utilized in this study were ATR-FTIR spectroscopy which infers the solute concentration and FBRM which provides crystal size information throughout the batch. Inferential modelling was used to construct a calibration curve to relate the FTIR spectra to the solute concentration, using procedures described elsewhere [158, 160]. FBRM measures the chord length distribution (CLD), which is not the same as the crystal size distribution (CSD) that appears in the models in the previous section.

The CSD can be computed from the CLD under certain assumptions [68, 135, 146, 154]. For some systems, the square-weighted chord length was found to be comparable to laser diffraction, sieving, and electrical sensing zone analysis over the range of $50 - 400\,\mu$m [59]. Although the aforementioned methods are able to estimate the CSD from CLD successfully for some systems, the theory behind these methods require many assumptions, including that the particles perfectly backscatter light at all angles and that shape of the crystals is known. Although these assumptions are true for many particulate systems (such as round polymer beads with a rough surface in water at low-to-moderate solids densities [68]), the assumptions

are not accurate for other particulate systems including the system studied here which has crystals with a similar refractive index as the solution (and hence poor backscattering properties). Due to the limited time and pharmaceutical quantity available in the early-stage of batch crystallization design, it is typically not possible to carry out the extensive studies to verify the assumptions and to determine the effects of non-ideality of the assumptions on the accuracy of the estimates of the CSD from the CLD. Furthermore, computing the CSD from the CLD when assumptions such as perfect laser backscattering do not hold is still an open problem [135, 161].

An alternative approach is to use the low-order moments of the CLD directly [54, 161] without first estimating the CSD from the CLD. This approach replaces the first-principles model for the CSD with a gray-box model for the CLD, in which the structure of the first-principles model for the low-order moments of the CSD is used to parametrize the low-order moments for the CLD [161]. The reasoning behind this particular gray-box model is that the mapping between the CLD and the CSD is static (most of the aforementioned mapping methods assume that the mapping is actually linear), so the low order moments of the CLD should follow the same dynamic trends as the low-order moments of the CSD. Due to the limitation of the FBRM precision, the zeroth moment was not used because FBRM would undercount the very small crystals. On the other hand, it is not advisable to use moments with order higher than two because higher order moments are sensitive to low-sampling statistics of the large crystals [54]. In this study, the first-order

moment was used. As with any model [5], this study assesses the applicability of this gray-box modelling by quantifying the accuracy of the kinetic parameters and the model's predictions.

The experiments are categorized into two sets, namely, α-seeded and β-seeded experiments. The seed crystal size distribution was approximated as a normal distribution

$$f_i\left(L,0\right) = f_{\mathrm{seed,i}}\left(L\right) = \frac{\lambda_i}{\sqrt{2\pi}\sigma_{\mathrm{seed},i}} \exp\left(-\frac{\left(L-\mu_{\mathrm{seed},i}\right)^2}{2\sigma_{\mathrm{seed},i}^2}\right), \tag{3.30}$$

with the parameters (λ_i, $\sigma_{\mathrm{seed},i}$, and $\mu_{\mathrm{seed},i}$) in Table 3.4. The time series for the temperature, first-order moment of the i-form crystals, and solute concentration for all experiments are shown by the solid lines in Figures 3.2-3.7. For all the β-seeded experiments, there is no apparent formation of α-form crystals at the end of all batches (Table 3.4).[§] As a result, the kinetic parameters for β-form crystals were independently obtained from the β-seeded experiments, except for $k_{b\beta,1}$, which accounts for the nucleation of β- from α-form crystals. One α-seeded experiment was operated at a high enough temperature that a measurable quantity of β-form crystals nucleated and grew (Experiment 3 in Table 3.4), so there would be enough information content in the data for $k_{b\beta,1}$ to be estimated. This experimental design enabled the kinetic parameters for β-form crystals to be obtained before determining the kinetic parameters for α-form crystals.

[§]Samples were taken at the end of all batches and XRD was used to determine the crystals form purity.

The nucleation and growth kinetics of α and β-form crystals have ten parameters to be estimated, four ($k_{b\alpha}$, $k_{g\alpha,0}$, g_α, $E_{g\alpha}$) corresponding to the kinetics of α-form crystals and six ($k_{b\beta,1}$, $k_{b\beta,2}$, $k_{g\beta,0}$, $k_{g\beta,2}$, g_β, $E_{g\beta}$) corresponding to the kinetics of β-form crystals. In relation to the notation defined in Section 3.3, the measured variables $\mathbf{y}$ and parameters of interest θ for each set of experiments are defined in Table 3.5, where σ_{ci}, $\sigma_{\mu_{i,1}}$, σ_{x_i} are the noise parameters for the i-form crystals. The prior distribution $\Pr(\theta)$ came from a preliminary parameter estimation that was carried out using maximum likelihood techniques as described in Miller and Rawlings [109], which resulted in a normal distribution for each parameter. These were modified for g_α and g_β according to Eq. (3.4) to limit their values between 1 and 2. The resulting marginal probability distributions of θ from α- and β-seeded experiments are in Figures 3.8 and 3.9, respectively. While some of the marginal probability distributions could be approximated by a normal distribution, others are not. These distributions can be directly inserted into those model predictive control and other control algorithms that have been designed to ensure robustness to stochastic parameter uncertainties [114]. The means, modes, and 95% credible intervals for the model parameters based on their marginal probability distributions are in Table 3.6. Figures 3.2-3.7 compare the temperature, first-order moment of the i-form crystals, and solute concentration trajectories obtained from experimental data and those predicted through simulation using the aforementioned mean values as the model parameters.

It is well-known that concentration data alone are not sufficient to character-

ize nucleation [109]. The small uncertainties in the nucleation kinetic parameters indicate that the first-order moment of the FBRM provided enough information to characterize the nucleation kinetics. The small range in the uncertainties for the activation energies indicates that the temperature range from 24 to 55°C in the experiments was large enough to enable activation energies to be estimated. The rather large uncertainty in $k_{g\alpha,0}$ is mainly due to the large correlation coefficient of 0.993 between $k_{g\alpha,0}$ and $E_{g\alpha}$, where a small change in $E_{g\alpha}$ necessitates a larger change in $k_{g\alpha,0}$ to ensure the resulting $k_{g\alpha}$ in Eq. (3.28) is of the same order of magnitude. Similar reasoning explains the large uncertainty in $k_{g\beta,0}$, with the correlation coefficient between $k_{g\beta,0}$ and $E_{g\beta}$ equal to 0.997. The growth exponent for the α-form is near 2, which indicates that the α-form growth rate is surface integration-limited, whereas that for the β-form is near 1, suggesting that the β-form growth rate is diffusion-limited. Unlike past studies that quantified uncertainties in the kinetic parameters for crystallization processes [109, 161], the analysis in this study explicitly takes into account hard theoretical bounds on the values for the parameters. In particular, the application of the linearized analyses used in past papers would have resulted in a confidence interval that included values of $g_\beta < 1$, whereas the Markov Chain simulation approach takes the lower bound of 1 into account during the statistical analysis (see Figure 3.9d).

To assess the predictive capability of the resulting model, another pair of experiments (i.e., one α- and one β-seeded experiment) were carried out with the seed distributions in Table 3.7. The trajectories of the temperature, first-order moment

of the i-form crystals, and solute concentration trajectories obtained from experimental data and those predicted through simulation are plotted in Figures 3.10 and 3.11. As can be seen from those figures, the predictive capability of the model is sufficiently accurate for use in process design and control. The solute concentration predicted by the model are quite close to the measured solute concentration in both validation experiments, with the differences between the predicted and experimental first-order moment being comparable to or smaller than the differences in the model and experimental first-order moments in the experiments used for parameter estimation (compare Figures 3.10 and 3.11 with Figures 3.2-3.7). The biases observed in the model predictions for the first-order moment of the i-form crystals could be due to the FBRM undercounting very small and large crystals, which would cause a different time-varying bias in different experiments.

Table 3.4: Seed crystal size distribution data and the purity of α-form crystals at the end of batch (x_α).

No.	Seed	Size $[\mu\mathrm{m}]$	Mass $[\mathrm{g/kg}]$	λ_i	$\sigma_{\mathrm{seed},i}$ $[\mathrm{m}] \times 10^6$	$\mu_{\mathrm{seed},i}$ $[\mathrm{m}] \times 10^6$	x_α
1	α	$180 - 250$	0.613	8.227×10^7	8.608	214.977	≈ 1.000
2	α	$75 - 180$	0.613	3.877×10^8	12.127	127.269	≈ 1.000
3	α	$75 - 180$	0.592	3.731×10^8	12.115	127.427	0.924
4	β	$40 - 270$	4.900	2.483×10^{10}	27.289	155.069	≈ 0.000
5	β	$40 - 270$	3.225	1.630×10^{10}	27.989	155.017	≈ 0.000
6	β	$40 - 270$	2.972	1.501×10^{10}	28.131	155.004	≈ 0.000

Table 3.5: Definition of measured variables $\mathbf{y}$ and interested parameters θ for α- and β-seeded experiments.

Seed	θ^T	$\mathbf{y}^T$
α	$[\ln(k_{b\alpha}), \ln(k_{g\alpha,0}), g_\alpha, \ln(E_{g\alpha}), \ln(k_{b\beta,1}), \ln(\sigma_{c\alpha}), \ln(\sigma_{\mu_{\alpha,1}}), \ln(\sigma_{x_\alpha})]$	$[C, \mu_{\alpha,1}, x_\alpha]$
β	$[\ln(k_{b\beta,2}), \ln(k_{g\beta,0}), \ln(k_{g\beta,2}), g_\beta, \ln(E_{g\beta}), \ln(\sigma_{c\beta}), \ln(\sigma_{\mu_{\beta,1}})]$	$[C, \mu_{\beta,1}]$

Table 3.6: The model parameters determined from parameter estimation.

Parameters	mean	mode	95% credible interval
$\ln(k_{b\alpha})$	17.233	17.213	17.083 to 17.377
$\ln(k_{g\alpha,0})$	1.878	1.778	0.801 to 2.912
g_α	1.859	1.860	1.775 to 1.944
$\ln(E_{g\alpha})$	10.671	10.671	10.612 to 10.725
$\ln(k_{b\beta,1})$	15.801	15.796	15.758 to 15.842
$\ln(k_{b\beta,2})$	20.000	20.000	19.961 to 20.036
$\ln(k_{g\beta,0})$	52.002	52.426	50.745 to 53.322
$\ln(k_{g\beta,2})$	-0.251	-0.251	-0.311 to -0.197
g_β	1.047	1.016	1.002 to 1.143
$\ln(E_{g\beta})$	12.078	12.076	12.060 to 12.097

Table 3.7: Seed crystal size distribution data and the purity of α-form crystals at the end of batch (x_α) for model validation.

No.	Seed	Size $[\mu\text{m}]$	Mass $[\text{g/kg}]$	λ_i	$\sigma_{\text{seed},i}$ $[\text{m}] \times 10^6$	$\mu_{\text{seed},i}$ $[\text{m}] \times 10^6$	x_α
V1	α	$75 - 180$	0.613	3.877×10^8	12.127	127.269	≈ 1.000
V2	β	$40 - 270$	3.060	1.547×10^{10}	28.081	154.978	≈ 0.000

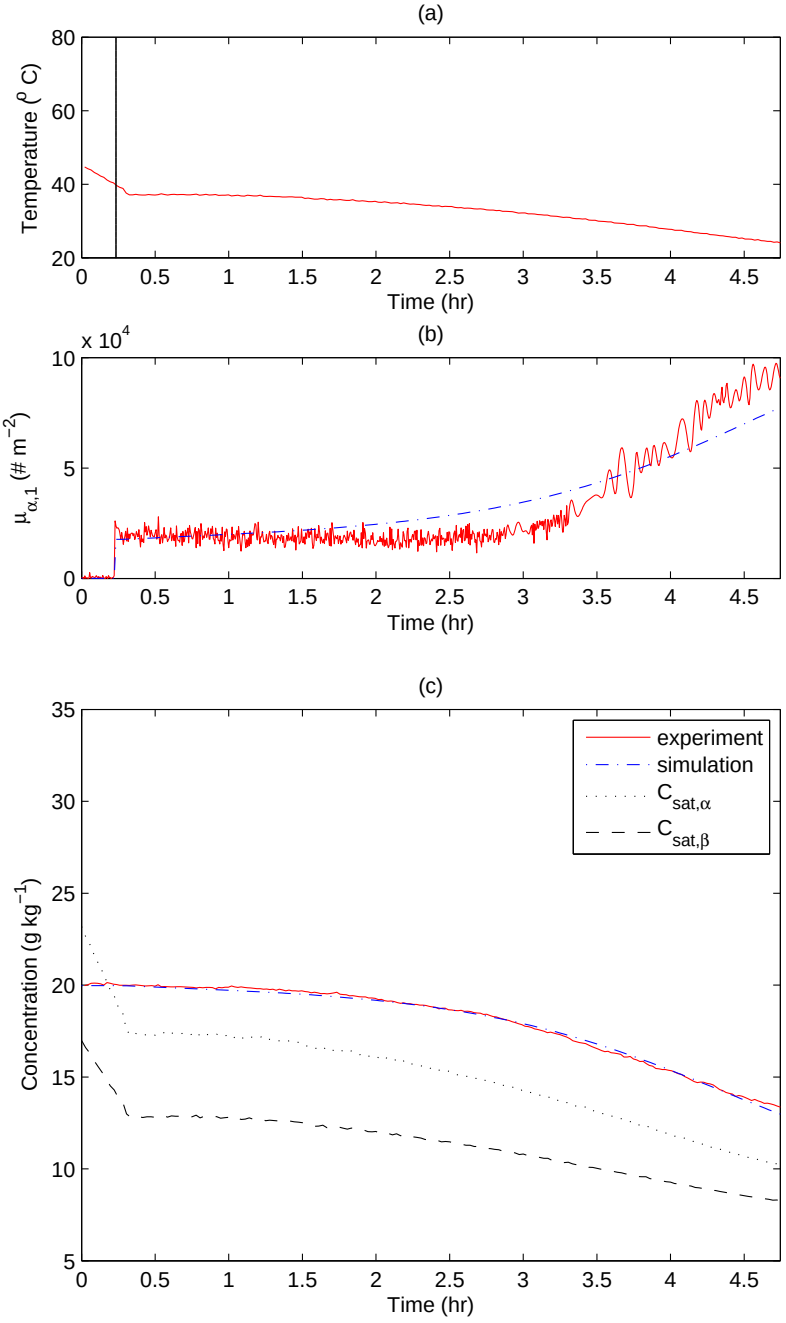


Figure 3.2: Experimental and model trajectories for (a) temperature, (b) the first-order moment of the α-form crystals, and (c) solute concentration for Experiment 1 of Table 3.4. The vertical line in plot (a) shows the seeding time.

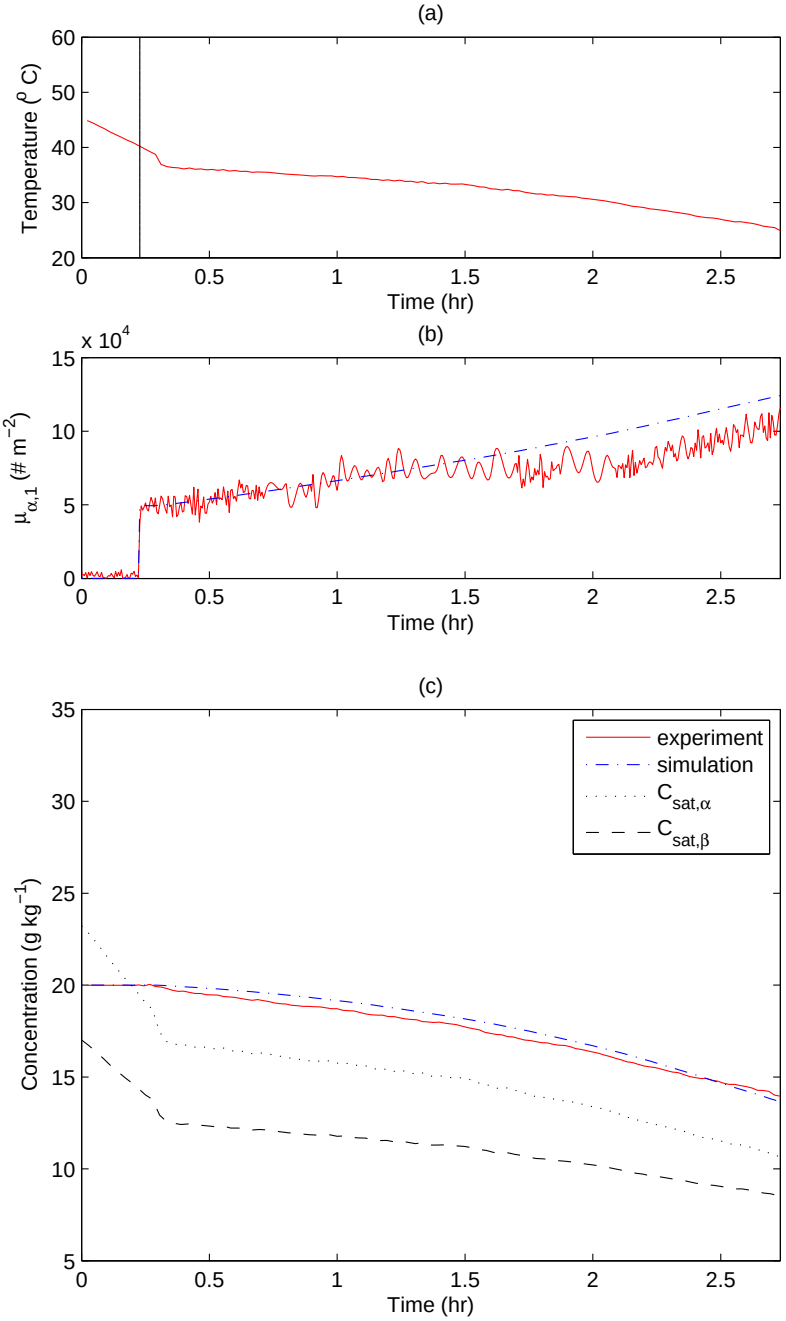


Figure 3.3: Experimental and model trajectories for (a) temperature, (b) the first-order moment of the α-form crystals, and (c) solute concentration for Experiment 2 of Table 3.4. The vertical line in plot (a) shows the seeding time.

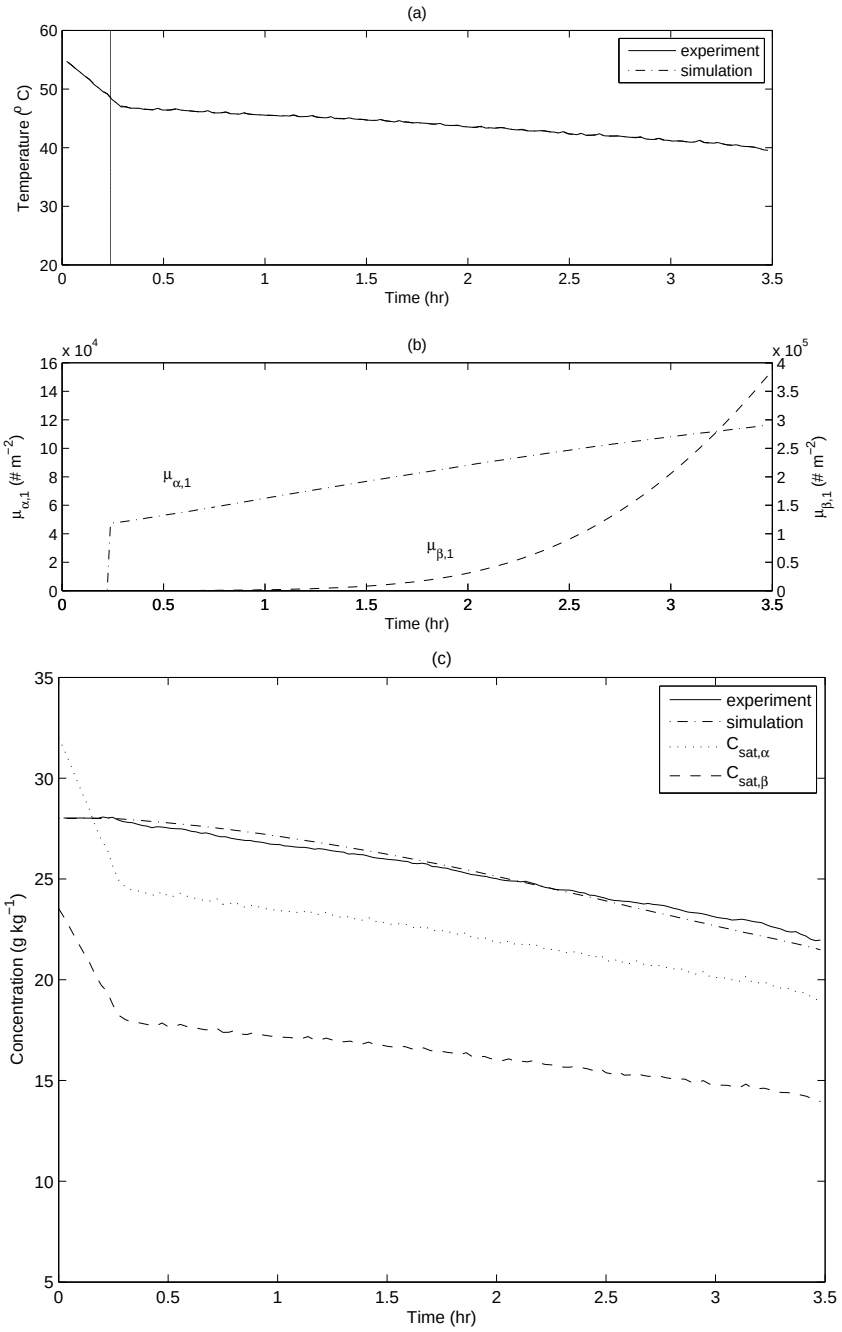


Figure 3.4: Experimental and model trajectories for (a) temperature and (b) the first-order moment of the α and β-form crystals, and (c) solute concentration for Experiment 3 of Table 4. The vertical line in plot (a) shows the seeding time. The experimental trajectory of the first-order moment is not plotted because the FBRM data was corrupted due to sensor fouling. Hence, the first-order moment from this experiment was not used in the parameter estimation.

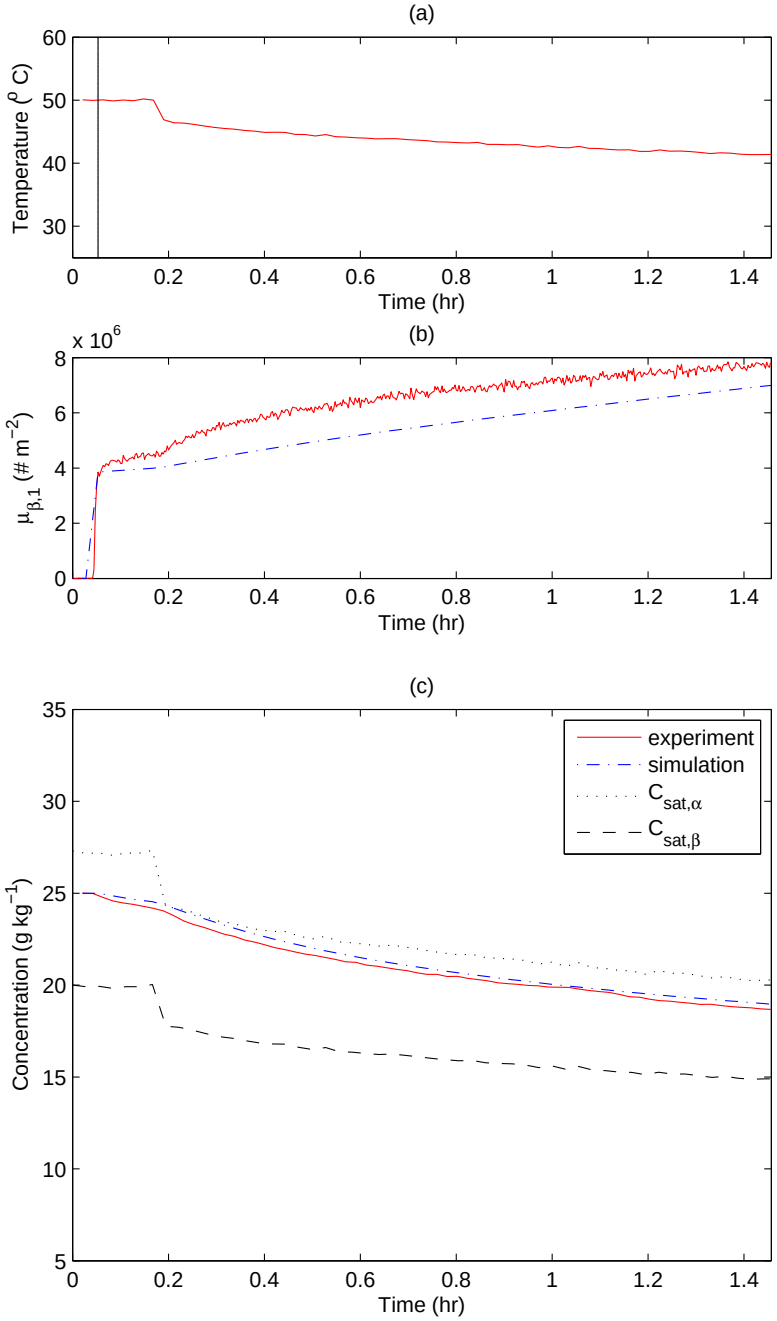


Figure 3.5: Experimental and model trajectories for (a) temperature, (b) the first-order moment of the β-form crystals, and (c) solute concentration for Experiment 4 of Table 3.4. The vertical line in plot (a) shows the seeding time.

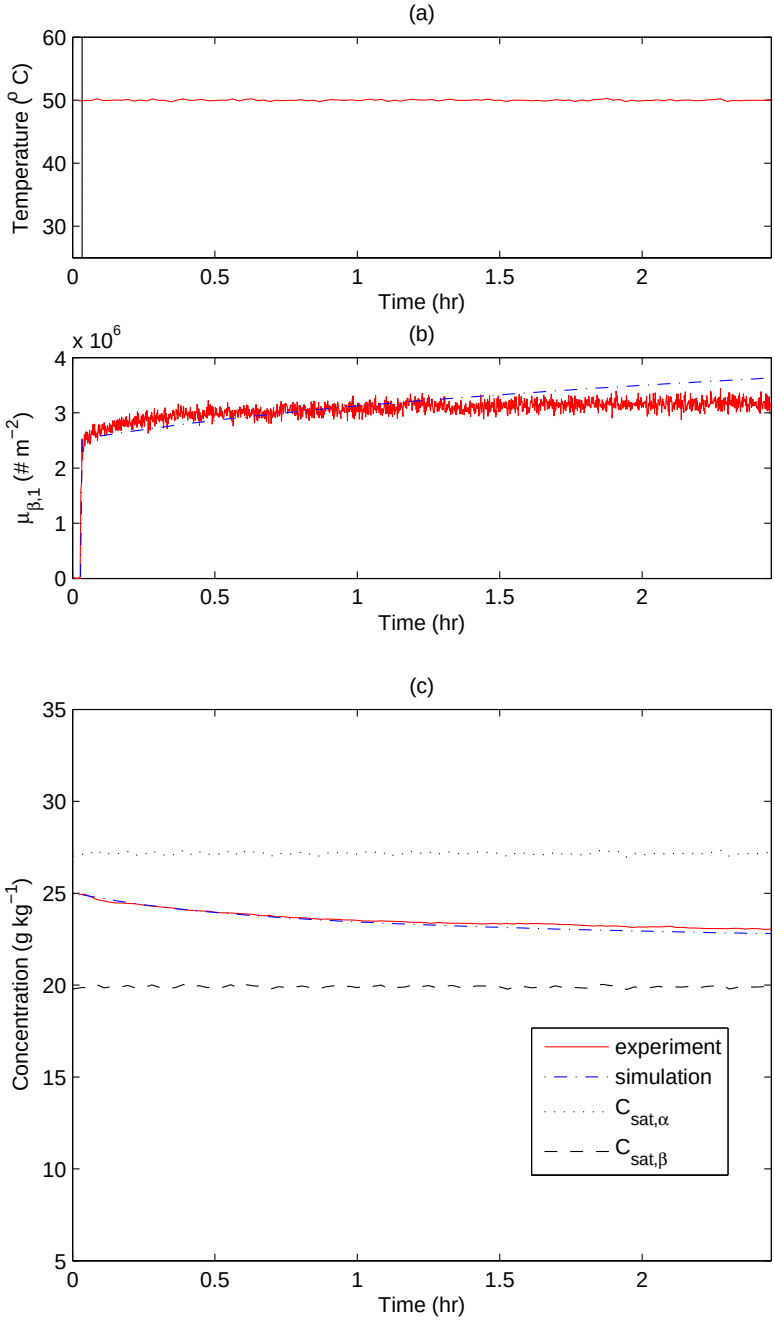


Figure 3.6: Experimental and model trajectories for (a) temperature, (b) the first-order moment of the β-form crystals, and (c) solute concentration for Experiment 5 of Table 3.4. The vertical line in plot (a) shows the seeding time.

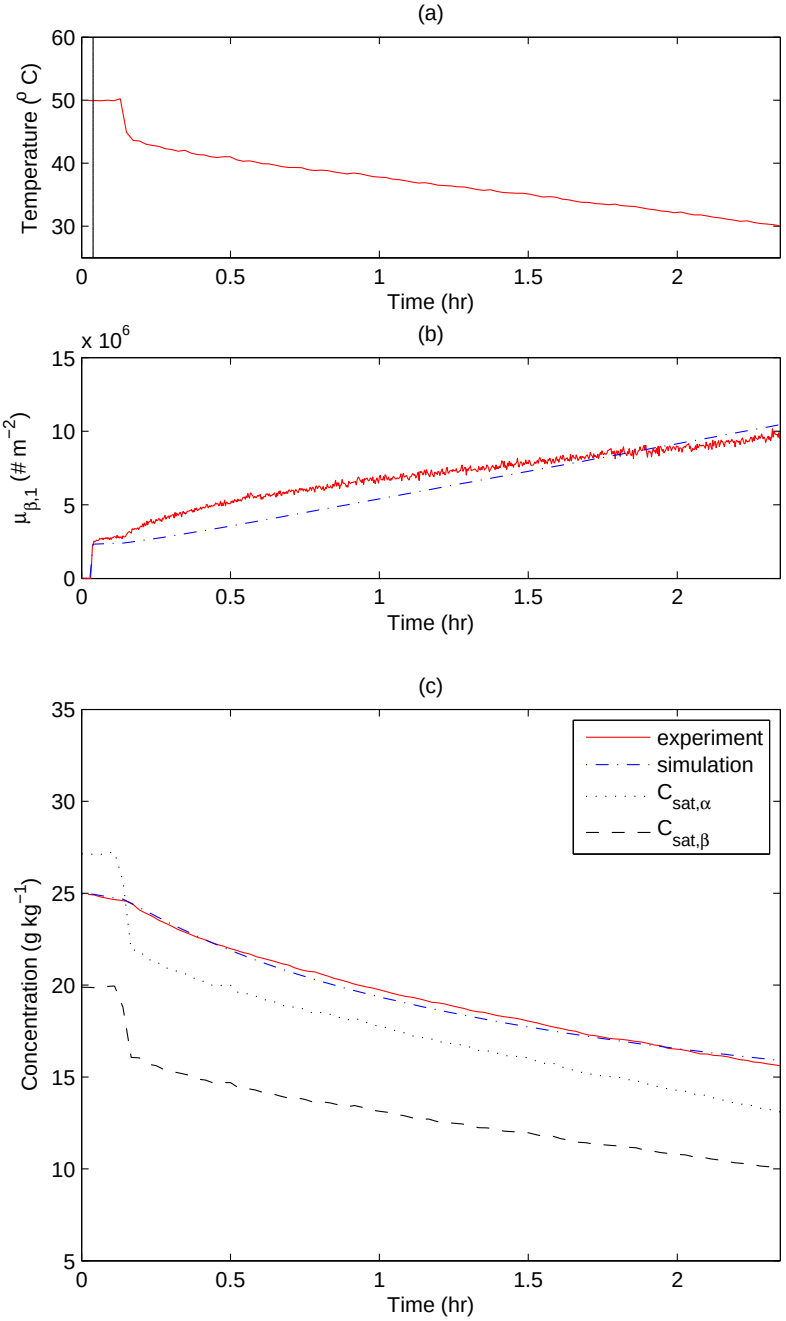


Figure 3.7: Experimental and model trajectories for (a) temperature, (b) the first-order moment of the β-form crystals, and (c) solute concentration for Experiment 6 of Table 3.4. The vertical line in plot (a) shows the seeding time.

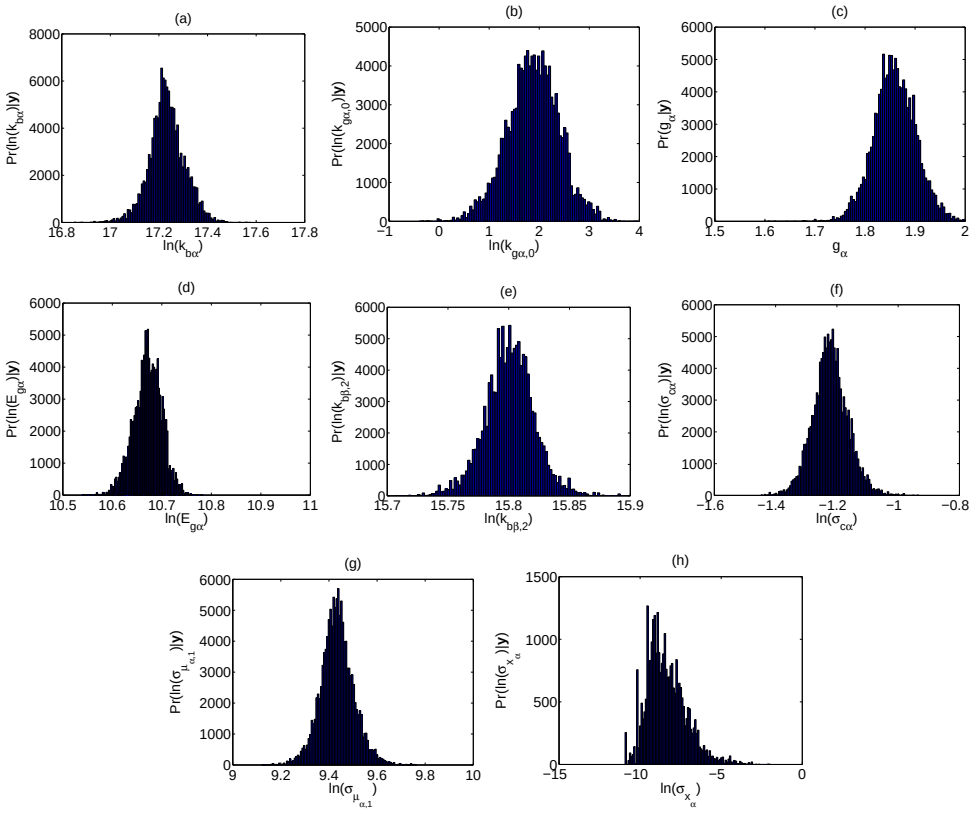


Figure 3.8: The marginal distributions of parameters θ obtained from α-seeded experiments (Table 3.5).

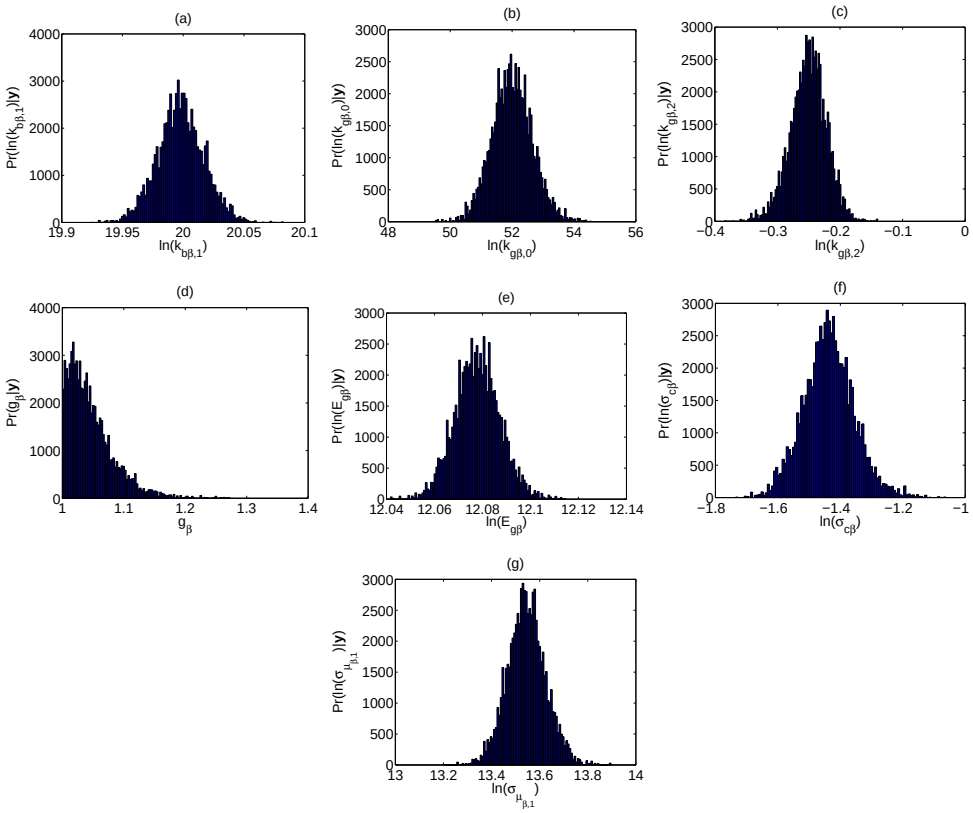


Figure 3.9: The marginal distributions of parameters θ obtained from β-seeded experiments (Table 3.5).

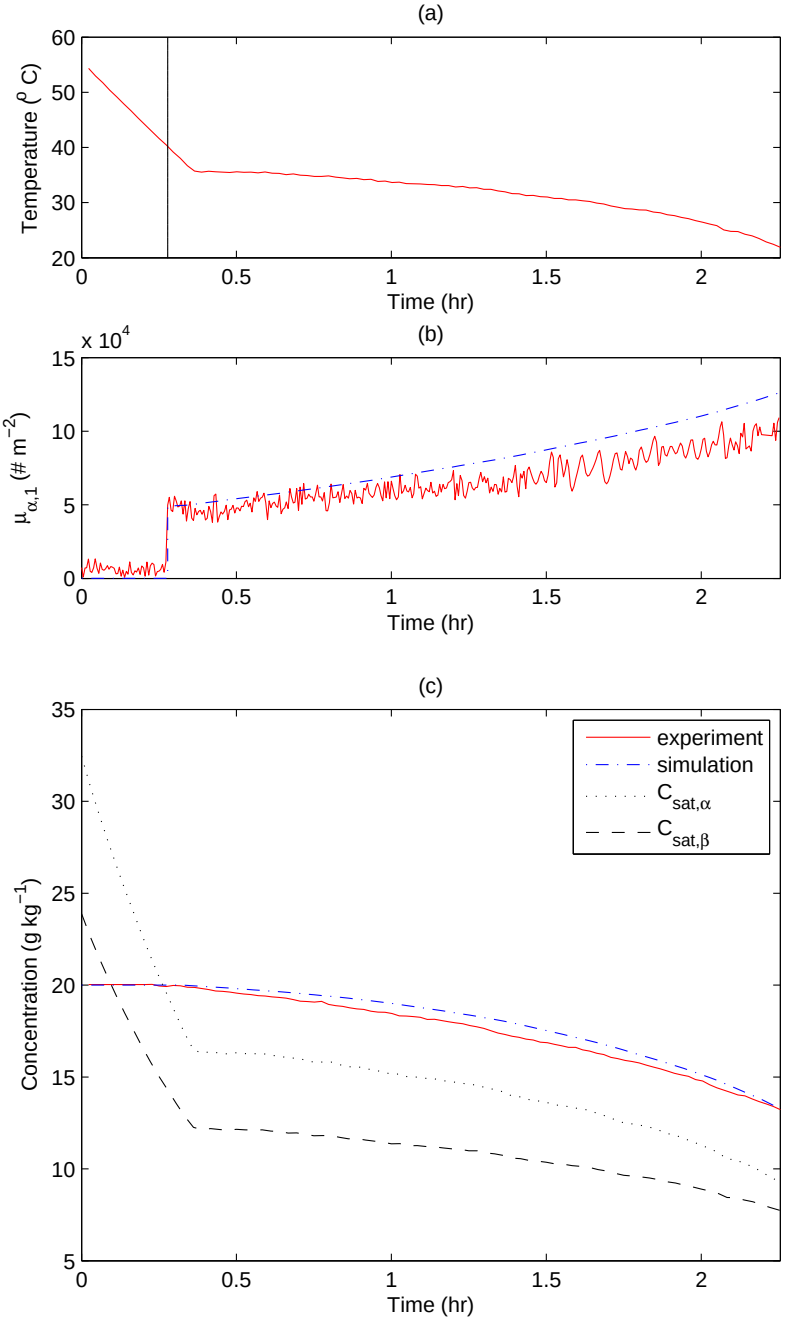


Figure 3.10: Experimental and predictive trajectories of (a) temperature, (b) the first-order moment of the α-form crystals, and (c) solute concentration for Experiment V1 of Table 3.7. The vertical line in plot (a) shows the seeding time.

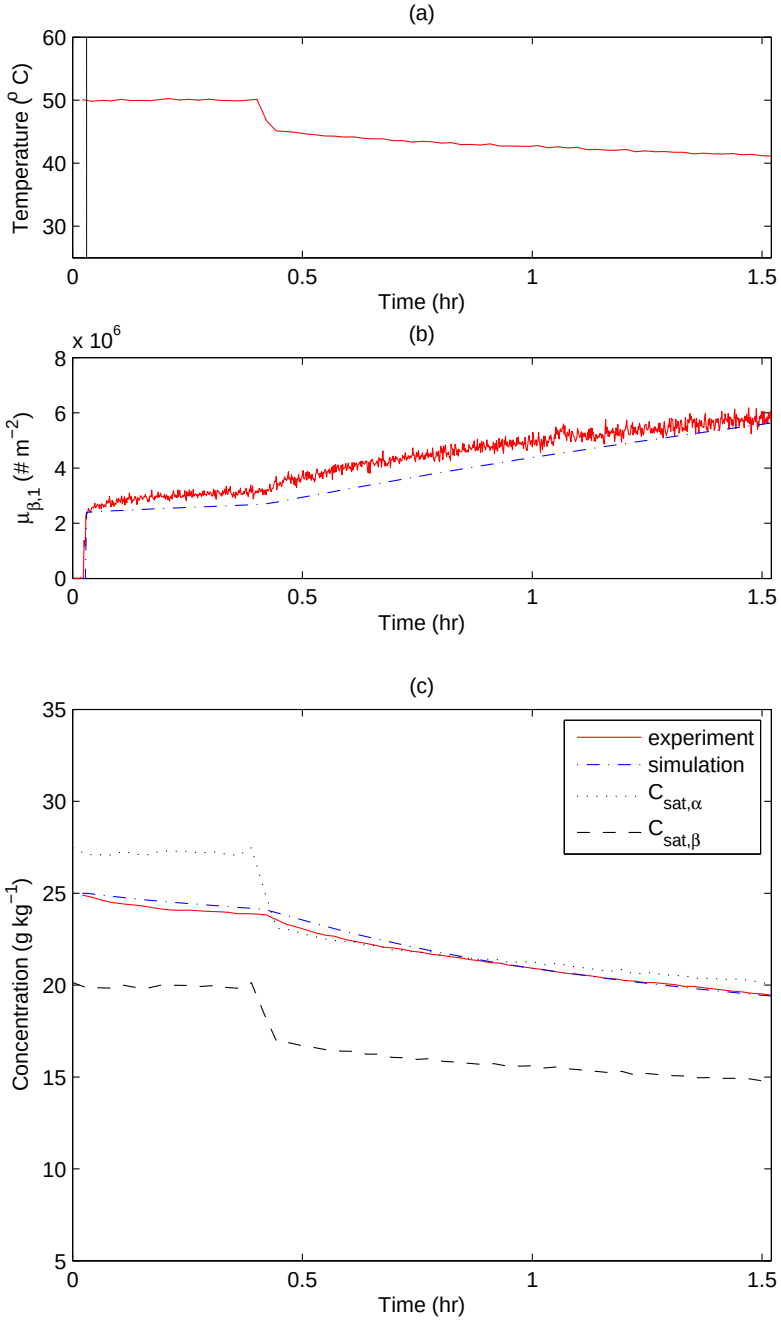


Figure 3.11: Experimental and predictive trajectories of (a) temperature, (b) the first-order moment of the β-form crystals, and (c) solute concentration for Experiment V2 of Table 3.7. The vertical line in plot (a) shows the seeding time.

3.5 Conclusions

A model of polymorphic crystallization of L-glutamic acid, which consist of α- and β-form crystallization, has been developed. The detailed kinetics model takes into account the temperature dependence of the crystals growth kinetic parameters, compared to past studies on the modelling of L-glutamic acid crystallization [115, 139]. In addition to providing point estimates of the kinetic parameters, a Bayesian inference approach is used to determine a detailed marginal probability distribution for each parameter. The marginal probability distributions of the parameters can give practitioners insight regarding the parameter uncertainties and are of significant value to develop robust control strategies for the crystallization process [114].

Although this study considers a specific polymorphic crystallization, the same parameter estimation method can be applied for crystallizations in which many nucleation and growth rates occur simultaneously, or when there are no prior literature data or estimates for the model parameters. The details of the nucleation and growth rate expressions may be different, depending on the particular solute-solvent system. With multiple polymorphs in the crystallizer, improved parameter estimates would be obtained by including polymorph ratio measurements obtained from in-situ Raman spectroscopy in Eq. (3.3) [150].

Chapter 4

High-order Simulation of Polymorphic Crystallization

4.1 Introduction

Numerical simulations for polymorphic crystallizations enable the investigation of the effects of various operating conditions and can be used for optimal design and control [64, 130, 139]. Solving population balance equations is particularly challenging when the partial differential equations (PDEs) are hyperbolic with sharp gradients or discontinuities in the distribution [148]. Standard first-order methods require a very small grid size in order to reduce the numerical diffusion (i.e., smearing), whereas standard higher order methods introduce numerical dispersion (i.e., spurious oscillations), which usually results in a crystal size distribution with negative values. Efficient and sufficiently accurate computational methods for simu-

lating the population balance equations are required to ensure the behaviour of the numerical solution is determined by the assumed physical principles and not by the chosen numerical method.

There have been many papers on the numerical solution of population balance models. The method of moments approximates the distribution by its moments [69], which under certain conditions, converts the hyperbolic PDEs into a small number of ordinary differential equations (ODEs) that describe characteristics of the distribution. The method of moments does not apply to population balance equations (PBEs) which do not satisfy moment closure conditions. The method of weighted residuals approximates the size distribution by a linear combination of basis functions [147], which results in a system of ODEs. For most practical crystallizations, a large number of basis functions is needed to approximate the distribution, which results in high computational cost. The Monte Carlo method tracks individual particles, each of which exhibits stochastical behaviour according to a probabilistic model [15, 58, 123]. This approach is too computationally expensive for most industrial crystallizations. Another problem-specific numerical method for solving population balance equations is the method of characteristics [83, 121]. This method solves each population balance equation by finding curves in the characteristic size-time plane that reduce the equation to an ODE. While the method is highly efficient when the kinetics are simple, the approach does not generalize to more complex kinetics. Most publications on numerical methods for solving PBEs involve various types of discretizations and go by a variety

of names including “method of classes” and “discretized population balance equations” [65, 82, 100, 101, 118]. In recent years there have been several efforts to reduce the numerical diffusion and numerical dispersion for distributions which contain sharp gradients or discontinuities, which is common in batch crystallizations. High resolution finite volume methods (FVMs) popular in astrophysics and gas dynamics [56, 92, 93, 116, 153] were extended to the application of multidimensional population balance equations [52, 104, 119, 120, 166]. A typical implementation applies a first-order method near discontinuities or sharp gradients and a second-order method everywhere else, which results in less numerical dispersion than the second-order method and less numerical diffusion than the first-order method [52].

This study considers a class of numerical algorithms known as weighted essentially non-oscillatory (WENO) methods which were developed for especially accurate simulation of shock waves and provide much higher order accuracy than the previously considered methods for solving population balance models (PBMs). Three WENO methods are considered: Liu et al’s version of WENO (LOCWENO) [99], Jiang and Shu’s version of WENO with Henrick mapping (JSHWENO) [61, 72], and the weighted power ENO method (Wpower-ENO) [141]. These WENO methods are compared to the high resolution (HR) finite volume method and a second-order finite difference (FD2) method, for polymorphic crystallization of L-glutamic acid under conditions in which the distribution contains sharp gradients. In the next section, the five numerical methods are discussed and followed by the discussion of simulation results. Then, conclusions are provided.

4.2 Numerical methods

In order to monitor the CSD of both seeded and nucleated crystals, it is convenient to represent Eq. (3.19) as:

$$\frac{\partial f_{\mathrm{seed},i}}{\partial t} + \frac{\partial \left(G_i f_{\mathrm{seed},i}\right)}{\partial L} = 0\,, \tag{4.1}$$

$$\frac{\partial f_{\mathrm{nucl},i}}{\partial t} + \frac{\partial \left(G_i f_{\mathrm{nucl},i}\right)}{\partial L} = B_i \delta\left(L - L_0\right)\,, \tag{4.2}$$

where $f_{\mathrm{seed},i}$ and $f_{\mathrm{nucl},i}$ are the crystal size distributions of the i-form crystals (i.e. α- or β-form crystals) obtained from seed crystals and nucleated crystals $[\#/\mathrm{m}^4]$, respectively.

The numerical methods described here differ in terms of their discretization along the crystal size dimension (L), each of which produces a system of ODEs describing the time evolution of the crystal size distribution at the chosen discretized points L_k [138]. To provide a fair basis for comparison, the implementation of all of the methods integrated the ODEs using a fourth-order orthogonal Runge-Kutta Chebyshev method [1], which is a class of explicit RungeKutta methods with extended stability domains along the negative real axis. The stability properties of this method make it suitable for stiff problems.

It is advantageous for a numerical method to be *conservative*, that is, to ensure that a quantity remains conserved by calculating a single flux which describes the flow of that quantity between neighbouring cells [61, 92]. Although flux conservative schemes are normally formulated using finite volumes, a finite difference

scheme is utilized here based on the approach described in Shu [143]. The inhomogeneous PDEs with source terms (4.2) were converted into homogeneous PDEs with boundary conditions:

$$\begin{aligned} \frac{\partial f_{\mathrm{nucl},i}}{\partial t} + \frac{\partial \left(G_i f_{\mathrm{nucl},i}\right)}{\partial L} &= 0\,, \qquad (4.3) \\ f_{\mathrm{nucl},i}\left(L_0, t\right) &= \frac{B_i}{G_i}\,. \qquad (4.4) \end{aligned}$$

To simplify notation, Eqs. (4.1) and (4.3) are written in the same form

$$\begin{aligned} \frac{\partial u}{\partial t} + \frac{\partial p}{\partial L} &= 0\,, \\ \frac{\partial u}{\partial t} &= -\frac{\partial p}{\partial L}\,, \end{aligned} \qquad (4.5)$$

where u is $f_{\mathrm{seed},i}$ or $f_{\mathrm{nucl},i}$ and p is $G_i f_{\mathrm{seed},i}$ or $G_i f_{\mathrm{nucl},i}$. Equation (4.5) is discretized in the L domain with uniform intervals of size ΔL, $L_k = k\Delta L$ indicates the crystal size at node k, and $I_k = \left[L_{k-1/2}, L_{k+1/2}\right]$ is the kth cell. The conservative approximation to the spatial derivative is used:

$$\frac{du_k(t)}{dt} = -\frac{1}{\Delta L}\left(\hat{p}_{k+1/2} - \hat{p}_{k-1/2}\right), \qquad (4.6)$$

where u_k is the value of u at L_k and the numerical flux $\hat{p}_{k+1/2}$ approximates $h_{k+1/2} = h(L_{k+1/2})$ with $h(L)$ implicitly defined by [144]

$$p\left(u\left(L\right)\right) = \frac{1}{\Delta L}\int_{L-\Delta L/2}^{L+\Delta L/2} h\left(\xi\right) d\xi\,. \qquad (4.7)$$

For stability, it is important that upwinding is used in constructing the numerical flux $\hat{p}_{k+1/2}$. One way is to compute the Roe speed to determine the direction of the wind:

$$\bar{a}_{k+1/2} = \frac{p_{k+1} - p_k}{u_{k+1} - u_k}\,, \tag{4.8}$$

where p_k is the value of p at L_k.

In the context of process model, the Roe speed is

$$\bar{a}_{k+1/2} \approx G_i\,, \tag{4.9}$$

and

- if $G_i \geq 0$ then the wind blows from the left to the right and the numerical fluxes $\hat{p}_{k+1/2}$ and $\hat{p}_{k-1/2}$ are approximated by $p^-_{k+1/2}$ and $p^-_{k-1/2}$, respectively.
- if $G_i < 0$ then the wind blows from the right to the left and the numerical fluxes $\hat{p}_{k+1/2}$ and $\hat{p}_{k-1/2}$ are approximated by $p^+_{k+1/2}$ and $p^+_{k-1/2}$, respectively.

The difference between the values with superscript $\pm$ at the same location $L_{k+1/2}$ is due to the possibility of different stencils for cell I_k and for cell I_{k+1}, that is, $p^-_{k+1/2}$ is due to the stencil for cell I_k and $p^+_{k+1/2}$ is due to the stencil for cell I_{k+1} (see Figure 4.1). In the next sections, five reconstruction procedures are described to obtain $p^-_{k+1/2}$ and $p^+_{k-1/2}$ only, as $p^+_{k+1/2}$ can be readily derived from $p^+_{k'-1/2}$ for cell $I_{k'} = I_{k+1}$ and $p^-_{k-1/2}$ can be derived from $p^-_{k'+1/2}$ for cell $I_{k'} = I_{k-1}$.

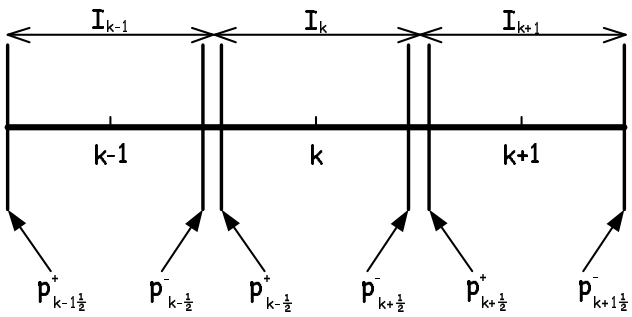


Figure 4.1: Computational cells.

4.2.1 WENO variants

All WENO methods discussed here are the derivatives of the original essentially non-oscillatory (ENO) method developed by Harten et al. [57] in 1987. This paper was the first to obtain a self similar (i.e., no mesh size-dependent parameter), uniformly high order accurate, yet essentially non-oscillatory interpolation (i.e., the magnitude of the oscillations decays as $O(\Delta x^r)$ where r is the order of accuracy) for piecewise smooth functions. ENO methods are especially suitable for problems containing both shocks and complicated smooth flow structures, such as those occurring in shock interactions with a turbulent flow and shock interaction with vortices. To improve the ENO method and further expand its applications, ENO methods based on point values and total diminishing variation (TVD) Runge-Kutta time discretizations were developed, which can reduce computational costs significantly for multiple space dimensions [143, 144]. Then biasing during selection of the stencil was proposed for enhancing stability and accuracy [39, 145]. Later,

WENO methods were developed, using a convex combination of all candidate stencils instead of just one as in the original ENO [61, 72, 99, 141].

WENO methods improve the accuracy of the original ENO method to the optimal order in smooth regions while maintaining the essentially non-oscillatory property near discontinuities. Liu et al [99] converted the rth order ENO method into an $(r+1)$th order WENO method with a cell average approach. Based on the pointwise finite difference ENO method [143, 144] and a new smoothness indicator, the WENO method by Jiang and Shu [72] can achieve the optimal $(2r-1)$th order accuracy. Jiang and Shu's WENO version was later modified by adding a mapping function for the original nonlinear weight which improves accuracy near smooth extrema [61]. Serna and Marquina [141] improved the behavior of Jiang and Shu's WENO method by introducing the powereno$_3$ or powermod$_3$ limiter, resulting in an $(2r-1)$th order weighted power ENO method. The powereno$_3$ or powermod$_3$ limiter substantially reduces smearing near discontinuities and results in better resolution of corners and local extrema.

All WENO methods adopt the following idea. Denote the r candidate stencils by

$$S_m = (L_{k+m-r+1}, L_{k+m-r+2}, \ldots, L_{k+m}), \quad m = 0, 1, \ldots, r-1, \tag{4.10}$$

whose corresponding rth order ENO approximation of the flux $h_{k+1/2}$ is

$$p^-_{k+1/2} = q^r_m(p_{k+m-r+1}, \ldots, p_{k+m})|_{L=L_{k+1/2}} . \tag{4.11}$$

Using the smoothest stencil among the r candidates for the approximation of $h_{k+1/2}$ is desirable near discontinuities to avoid introducing aphysical oscillations. All of the stencils are smooth in regions where the solution is smooth, in which case it is better to combine the results of multiple stencils together to produce a higher order (higher than rth order, the order of the original ENO method) approximation to the flux $h_{k+1/2}$ [72]. WENO methods assign a weight ω_m to each candidate stencil S_m to obtain the combined approximation of $h_{k+1/2}$ as

$$p^-_{k+1/2} = \sum_{m=0}^{r-1} \omega_m q^r_m(p_{k+m-r+1}, \ldots, p_{k+m})|_{L=L_{k+1/2}} . \tag{4.12}$$

To achieve the essentially non-oscillatory property, the weights adapt to the relative smoothness of p on each candidate stencil such that any discontinuous stencil is effectively assigned a zero weight. In smooth regions the weights are adjusted such that the resulting approximation gives an order of accuracy higher than r. The differences between WENO methods lie on the method for selecting the weights ω_m and the flux approximations $q^r_m(p_{k+m-r+1}, \ldots, p_{k+m})$. The subsequent WENO methods have $r = 3$ with the flux approximations $q^r_m(p_{k+m-r+1}, \ldots, p_{k+m})$ constructed based on quadratic polynomials.

Liu et al.'s WENO (LOCWENO) method

The flux approximations and weights for the fourth-order accurate LOCWENO method are [99]

$$\begin{aligned} q_m^3(p_{k+m-2}, p_{k+m-1}, p_{k+m}) &= \frac{p_{k+m} - 2p_{k+m-1} + p_{k+m-2}}{2\Delta L}\left(L - L_{k+m-1}\right)^2 \\ &\quad + \frac{p_{k+m} - p_{k+m-2}}{2\Delta L}\left(L - L_{k+m-1}\right) \\ &\quad + p_{k+m-1} - \frac{p_{k+m} - 2p_{k+m-1} + p_{k+m-2}}{24} \end{aligned} \tag{4.13}$$

and

$$\omega_m = \frac{\lambda_m}{\sum_{j=0}^{2} \lambda_j}\,, \tag{4.14}$$

where

$$\lambda_m = \begin{cases} \frac{d_m}{(IS_m + \epsilon)^h} & \text{for } p_{k+1/2}^- \\ \frac{d_{2-m}}{(IS_m + \epsilon)^h} & \text{for } p_{k-1/2}^+ \,. \end{cases} \tag{4.15}$$

The values of h and d_m in Eq. (4.15) are in Table 4.1 and ϵ is a small number to avoid division by zero (i.e., $\epsilon = 10^{-4}$ was used in this study). The IS_m are

Table 4.1: Values of h and d_m for LOCWENO, JSHWENO, and WPower-ENO methods.

	LOCWENO	JSHWENO	WPower-ENO
h	3	2	2
d_0	1/12	1/10	1/5
d_1	1/2	3/5	1/5
d_2	1/4	3/10	2/5

smoothness indicators given by

$$\begin{aligned} IS_m &= \frac{\left(p_{k+m-1}-p_{k+m-2}\right)^2+\left(p_{k+m}-p_{k+m-1}\right)^2}{2} \\ &\quad +\left(p_{k+m}-2p_{k+m-1}+p_{k+m-2}\right)^2 . \end{aligned} \tag{4.16}$$

Jiang and Shu's WENO method with Henrick mapping (JSHWENO)

The Jiang and Shu's WENO method used here is based on the quadratic polynomial instead of the original linear approximation. With

$$d_{k+1/2} = p_{k+1}-p_k , \tag{4.17}$$

$$d_k = \frac{d_{k+1/2}+d_{k-1/2}}{2} , \tag{4.18}$$

$$D_k = d_{k+1/2}-d_{k-1/2} , \tag{4.19}$$

the flux approximations are

$$\begin{aligned} q_0^3\left(p_{k-2},p_{k-1},p_k\right) &= p_k-\frac{D_{k-1}}{24}+ \\ &\quad \frac{(L-L_k)}{\Delta L}\left[d_{k-1/2}+\frac{D_{k-1}}{2}+\frac{D_{k-1}}{2}\left(\frac{L-L_k}{\Delta L}\right)\right], \end{aligned} \tag{4.20}$$

$$q_1^3\left(p_{k-1},p_k,p_{k+1}\right) = p_k - \frac{D_k}{24} + \frac{(L-L_k)}{\Delta L}\left[d_k + \frac{D_k}{2}\left(\frac{L-L_k}{\Delta L}\right)\right], \tag{4.21}$$

$$q_2^3\left(p_k,p_{k+1},p_{k+2}\right) = p_k - \frac{D_{k+1}}{24} + \frac{(L-L_k)}{\Delta L}\left[d_{k+1/2} - \frac{D_{k+1}}{2} + \frac{D_{k+1}}{2}\left(\frac{L-L_k}{\Delta L}\right)\right], \tag{4.22}$$

and the weights are [72]

$$\omega_m^{JS} = \frac{\lambda_m}{\sum_{j=0}^{2}\lambda_j}, \tag{4.23}$$

where λ_m is defined in Eq. (4.15) and the values of h and d_m are in Table 4.1. The smoothness indicators IS_m are

$$IS_0 = \frac{13}{12}\left(p_{k-2}-2p_{k-1}+p_k\right)^2 + \frac{1}{4}\left(p_{k-2}-4p_{k-1}+3p_k\right)^2, \tag{4.24}$$

$$IS_1 = \frac{13}{12}\left(p_{k-1}-2p_k+p_{k+1}\right)^2 + \frac{1}{4}\left(p_{k-1}-p_{k+1}\right)^2, \tag{4.25}$$

$$IS_2 = \frac{13}{12}\left(p_k-2p_{k+1}+p_{k+2}\right)^2 + \frac{1}{4}\left(3p_k-4p_{k+1}+p_{k+2}\right)^2. \tag{4.26}$$

The Henrick mapping [61]

$$g_m(\omega) = \frac{\omega(d_m + d_m^2 - 3d_m\omega + \omega^2)}{d_m^2 + (1-2d_m)\omega}. \tag{4.27}$$

is used to revise these weights to improve the accuracy near smooth extrema:

$$\omega_m^{HJS} = \frac{\lambda_m^*}{\sum_{j=0}^{2} \lambda_j^*} \tag{4.28}$$

with

$$\lambda_m^* = g_m(\omega_m^{JS}) \,, \tag{4.29}$$

to produce a fifth-order accurate method.

Weighted power ENO (Wpower-ENO) method

Using the definitions in Eqs. (4.17)-(4.19), the flux approximations for the Wpower-ENO method are

$$\begin{aligned} q_0^3(p_{k-2}, p_{k-1}, p_k) &= p_k - \frac{Pow_{k-1/2}}{24} + \\ &\quad \frac{(L - L_k)}{\Delta L} \left[d_{k-1/2} + \frac{Pow_{k-1/2}}{2} + \frac{Pow_{k-1/2}}{2} \left(\frac{L - L_k}{\Delta L} \right) \right] , \end{aligned} \tag{4.30}$$

$$q_1^3(p_{k-1}, p_k, p_{k+1}) = p_k - \frac{D_k}{24} + \frac{(L - L_k)}{\Delta L} \left[d_k + \frac{D_k}{2} \left(\frac{L - L_k}{\Delta L} \right) \right] , \tag{4.31}$$

$$\begin{aligned} q_2^3(p_k, p_{k+1}, p_{k+2}) &= p_k - \frac{Pow_{k+1/2}}{24} + \\ &\quad \frac{(L - L_k)}{\Delta L} \left[d_{k+1/2} - \frac{Pow_{k+1/2}}{2} + \frac{Pow_{k+1/2}}{2} \left(\frac{L - L_k}{\Delta L} \right) \right] , \end{aligned} \tag{4.32}$$

where

$$Pow_{k+1/2} = \mathrm{powereno}_3(D_k, D_{k+1})\,, \tag{4.33}$$

is the powereno limiter acting on $L = L_{k+1/2}$ where

$$\mathrm{powereno}_3(x,y) = \mathrm{minsign}(x,y)\cdot \mathrm{power}_3(|x|,|y|)\,, \tag{4.34}$$

$$\mathrm{minsign}(x,y) = \begin{cases} \mathrm{sign}(x) & \text{if } |x| \le |y|\,, \\ \mathrm{sign}(y) & \text{otherwise}\,, \end{cases} \tag{4.35}$$

$$\mathrm{power}_3(x,y) = \min(x,y)\frac{x^2+y^2+2[\max(x,y)]^2}{(x+y)^2}\,. \tag{4.36}$$

The weights ω_m and parameters λ_m are the same as for the LOCWENO method, except that the smoothness indicators

$$IS_0 = \frac{13}{12}\left(Pow_{k-1/2}\right)^2 + \frac{1}{4}\left(2p_k - 2p_{k-1} + Pow_{k-1/2}\right)^2\,, \tag{4.37}$$

$$IS_1 = \frac{13}{12}\left(p_{k-1} - 2p_k + p_{k+1}\right)^2 + \frac{1}{4}\left(p_{k-1} - p_{k+1}\right)^2\,, \tag{4.38}$$

$$IS_2 = \frac{13}{12}\left(Pow_{k+1/2}\right)^2 + \frac{1}{4}\left(2p_{k+1} - 2p_k - Pow_{k+1/2}\right)^2\,. \tag{4.39}$$

are used. This method is fifth-order accurate.

4.2.2 High resolution (HR) method

The popular high resolution method uses second-order discretization with a flux limiter to ensure non-oscillatory behaviour. For $G_i \ge 0$, a backward second-order

discretization is used:

$$p^-_{k+1/2} - p^-_{k-1/2} = \frac{1}{2}(3p_k - 4p_{k-1} + p_{k-2}) = \frac{1}{2}(3p_k - p_{k-1}) - \frac{1}{2}(3p_{k-1} - p_{k-2}) \tag{4.40}$$

or

$$p^-_{k+1/2} = \frac{1}{2}(3p_k - p_{k-1}) = p_k + \frac{1}{2}(p_k - p_{k-1})\,, \tag{4.41}$$

where the first term is first-order and the second term is commonly referred to as an "anti-diffusion term" because it reduces numerical diffusion. Applying a flux limiter on the anti-diffusion term yields

$$p^-_{k+1/2} = p_k + \frac{1}{2}\phi(w_k)(p_k - p_{k-1})\,, \tag{4.42}$$

where w_k is the upwinding ratio defined by

$$w_k = \frac{p_{k+1} - p_k}{p_k - p_{k-1}} \tag{4.43}$$

and $\phi(\cdot)$ is the flux limiter. In this study, the popular Van Leer flux limiter [162] was used:

$$\phi(w) = \frac{w + |w|}{1 + w}\,. \tag{4.44}$$

For $G_i < 0$, a forward second-order discretization is used:

$$p^+_{k+1/2} - p^+_{k-1/2} = \frac{1}{2}(-3p_k + 4p_{k+1} - p_{k+2}) = \frac{1}{2}(-p_{k+2} + 3p_{k+1}) - \frac{1}{2}(-p_{k+1} + 3p_k) \tag{4.45}$$

or

$$p^+_{k-1/2} = \frac{1}{2}(3p_k - p_{k+1}) = p_k - \frac{1}{2}(p_{k+1} - p_k) \tag{4.46}$$

Similar inclusion of a flux limiter to the anti-diffusion term gives

$$p^+_{k-1/2} = p_k - \frac{1}{2}\phi\left(\frac{1}{w_k}\right)(p_{k+1} - p_k)\,. \tag{4.47}$$

This high resolution method is second-order accurate in smooth regions, and first-order accurate near discontinuities.

4.2.3 The second-order finite difference (FD2) method

A second-order finite difference method with correct upwinding uses the fluxes $p^-_{k+1/2}$ and $p^+_{k-1/2}$ given by Eqs. (4.41) and (4.46), respectively.

4.3 Simulation results

The five numerical methods were applied to the L-glutamic acid polymorphic crystallization model. The initial seed distributions $f_{\text{seed},i}(L,0)$ for α- and β-forms are described by Gaussian distributions:

$$f_{\text{seed},i}(L,0) = \frac{\kappa_i}{\sqrt{2\pi}\sigma_{\text{seed},i}} \exp\left(-\frac{(L-\mu_{\text{seed},i})^2}{2\sigma_{\text{seed},i}^2}\right), \tag{4.48}$$

with the parameters in Table 4.2 selected so that the distributions would be sharp enough to challenge numerical methods. The temperature profile is in Figure 4.2 where the vertical solid line indicates the seeding time (i.e., at $t = 10$ min). Since an analytical solution is not available, the reference solutions for all CSDs were obtained by using WPower-ENO method with very fine resolution. All the computations were performed using Compaq Fortran 6.6 on a HP workstation XW6400 (Intel Xeon 5150 (2.66 GHz) and 2 GB of RAM).

In an unseeded crystallization the CSD profiles for all methods are nearly coincident with the reference profiles (see Figures 4.3 and 4.4), indicating that a conventional numerical method such as FD2 might suffice, which is consistent with expectations since no sharp gradients occur in these distributions. In the case of seeded crystallization (the usual case in practice), the differences in the CSD profiles between the WENO variants and their conventional counterparts are significant (see Figures 4.5 and 4.6). While the three WENO variants are nearly indistinguishable with the reference profiles, the HR and FD2 methods exhibit numerical diffusion

and do not resolve the peaks accurately. In addition, the FD2 method introduces a spurious oscillation (known as numerical dispersion) which can occur near sharp gradients with this method. The HR method does not produce spurious oscillation because the flux limiter detects the presence of sharp gradients and limits the size derivatives. The larger numerical errors in the CSD profiles obtained by HR and FD2 methods for the seeded α-form compared to the seeded β-form are associated with its sharper gradient.

The prediction errors were quantified in terms of the L_1 norm (E_{L_1}):

$$E_{L_1} = \frac{1}{2(N_{\text{grid,seed}} + N_{\text{grid,nucl}})} \times \sum_{i=\{\alpha,\beta\}} \left\{ \sum_{k=1}^{N_{\text{grid,seed}}} \left| f_{\text{seed},i,k} - f^{\text{ref}}_{\text{seed},i,k} \right| + \sum_{k=1}^{N_{\text{grid,nucl}}} \left| f_{\text{nucl},i,k} - f^{\text{ref}}_{\text{nucl},i,k} \right| \right\}, \tag{4.49}$$

where $f^{\text{ref}}_{\text{seed},i,k}$ and $f^{\text{ref}}_{\text{nucl},i,k}$ are the reference solutions for the seeded and nucleated crystals size distributions and $N_{\text{grid,seed}}$ and $N_{\text{grid,nucl}}$ are the number of grids used to discretize the size coordinate of the seeded and nucleated crystal size distributions, respectively. The error L_1 norms from the three WENO variants are much smaller in magnitude and grow much slower than those from the HR and FD2 methods (see Figure 4.7). In terms of the L_1 norm, the JSHWENO method gave the smallest numerical errors. Figure 4.8 indicates that the JSHWENO method gave smaller numerical errors for the full range of ΔL from 0.1 to 2.0 μm.

The JSHWENO method generally had lower CPU times than the WPower-ENO

method, but somewhat higher CPU times than the other methods for most values of ΔL (see Figure 4.9). To fairly compare the overall efficiency for these methods, the CPU time was compared for discretizations that produce the same error L_1 norm. From Figure 4.10 it is observed that, for any given error L_1 norm, the WENO variants used less or equal CPU time to the HR and FD2 methods, and hence the WENO variants were more efficient. The JSHWENO method was the most efficient for nearly all desired accuracy levels. Figure 4.11 shows the relative cost of the numerical methods with respect to the JSHWENO method. The HR method was more efficient than the FD2 method for nearly all desired accuracy levels, and was more efficient than the WPower-ENO method for some accuracy levels, but was not as efficient as the LOCWENO and JSHWENO methods. While the WENO methods are more complicated to implement, their efficiency is much better when sufficiently high accuracy in the size distribution is desired. Among the WENO variants, the performance of JSHWENO is followed by that of the LOCWENO method by a small margin, and then followed by that of the WPower-ENO method.

Another way to assess numerical methods is to compute the L_1 self-convergence order

$$O_{L_1} = \frac{\ln\left(\frac{E_{L_1}|_{2\Delta L}}{E_{L_1}|_{\Delta L}}\right)}{\ln 2}\,. \tag{4.50}$$

This metric provides information on the internal consistency of the numerical method and its intrinsic convergence [49]. The L_1 self-convergence order for all numerical

methods are in Table 4.3. For a linear model with a smooth solution, these values would correspond to the order of the truncation error for a given numerical method. This is not the case here due to the nonlinearity of the model and the sharp gradients in the distribution. On average, the JSHWENO method gives the best L_1 self-convergence order, followed by the WPower-ENO, LOCWENO, HR, and FD2 methods.

Table 4.2: Initial seed distribution parameters for α- and β-forms.

i	κ_i	$\sigma_{\text{seed},i}[\text{m}] \times 10^6$	$\mu_{\text{seed},i}[\text{m}] \times 10^6$
α	2×10^{10}	2.000	30.000
β	2×10^{10}	4.000	50.000

Table 4.3: L_1 self-convergence order (O_{L_1}) for the various numerical methods.

$\Delta L\,[\mu\text{m}]$	LOCWENO	JSHWENO	WPower-ENO	HR	FD2
0.1	1.62	1.30	1.53	1.67	1.81
0.2	2.43	1.45	1.97	1.66	1.78
0.3	2.41	2.40	2.12	1.65	1.57
0.4	2.26	2.84	2.28	1.57	1.37
0.5	2.27	2.87	2.55	1.50	1.20
0.6	2.46	2.98	2.79	1.44	1.07
0.7	2.41	3.12	2.86	1.34	0.98
0.8	2.24	3.06	2.68	1.25	0.88
0.9	2.11	3.06	2.60	1.20	0.82
1.0	1.94	2.86	2.39	1.09	0.75
average order	2.22	2.59	2.38	1.44	1.22

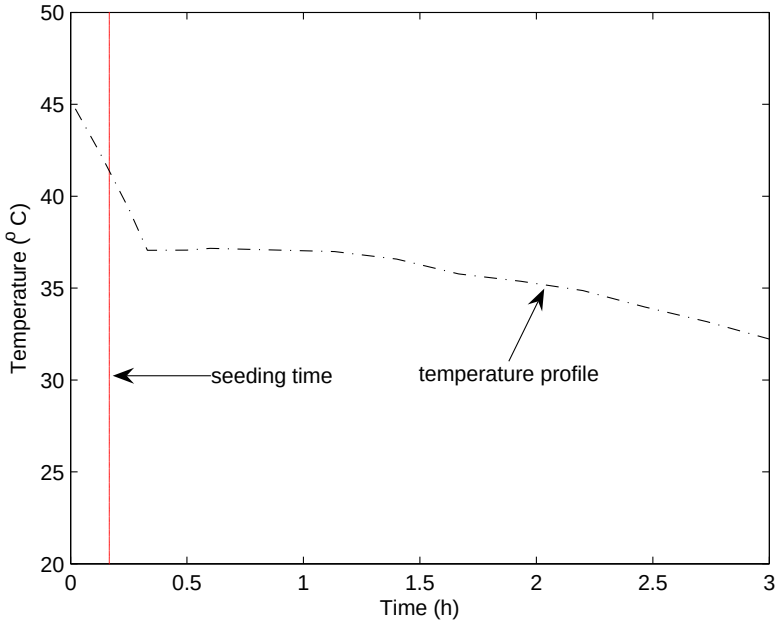


Figure 4.2: Temperature profile used in simulations.

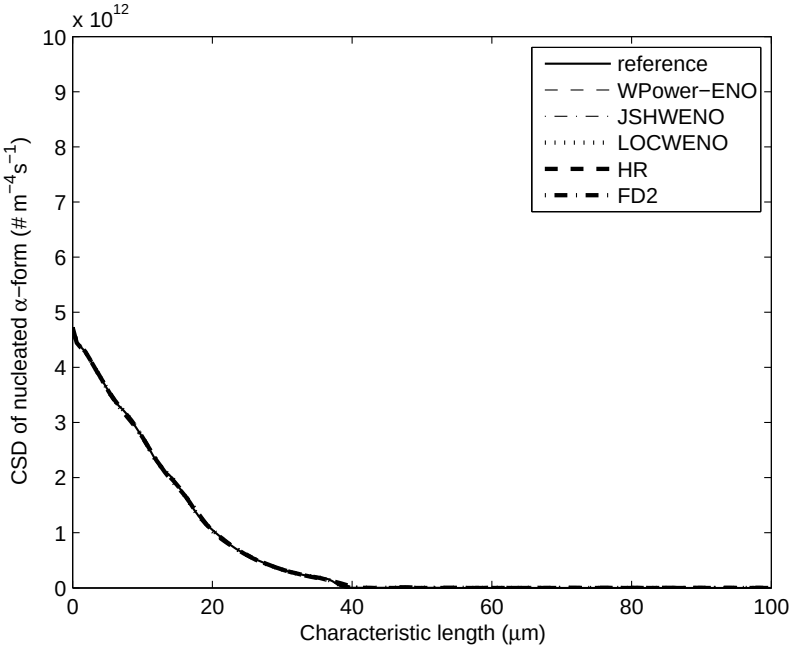


Figure 4.3: CSD of nucleated α form at the end of the batch for the various numerical methods ($\Delta L = 0.6\ \mu\mathrm{m}$).

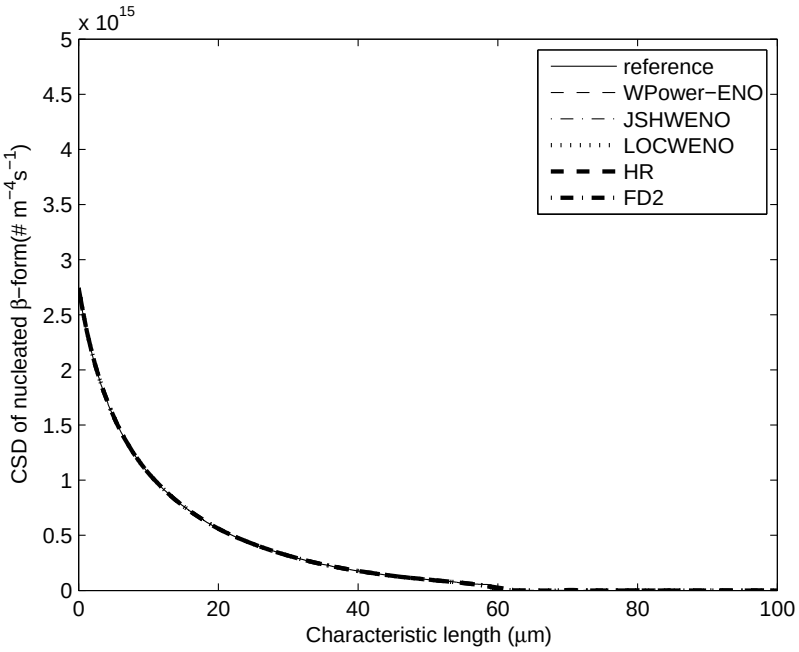


Figure 4.4: CSD of nucleated β form at the end of the batch for the various numerical methods ($\Delta L = 0.6\ \mu\mathrm{m}$).

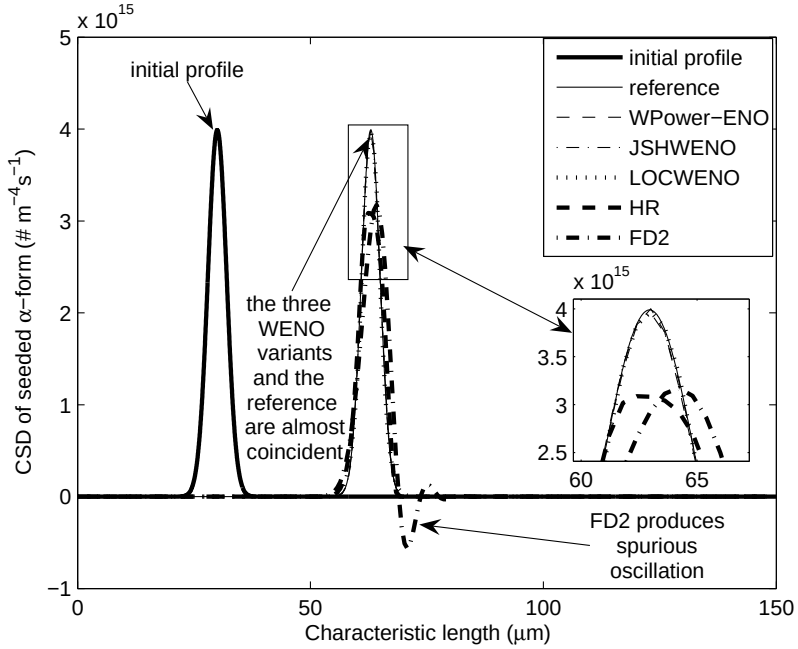


Figure 4.5: CSD of seeded α form at the end of the batch for the various numerical methods ($\Delta L = 0.6\ \mu\mathrm{m}$).

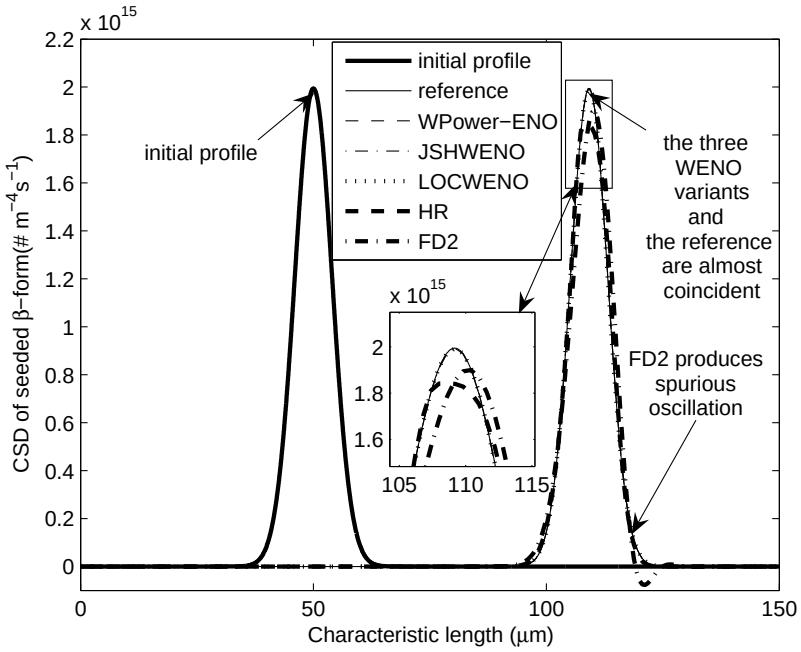


Figure 4.6: CSD of seeded β form at the end of the batch for the various numerical methods ($\Delta L = 0.6\ \mu\mathrm{m}$).

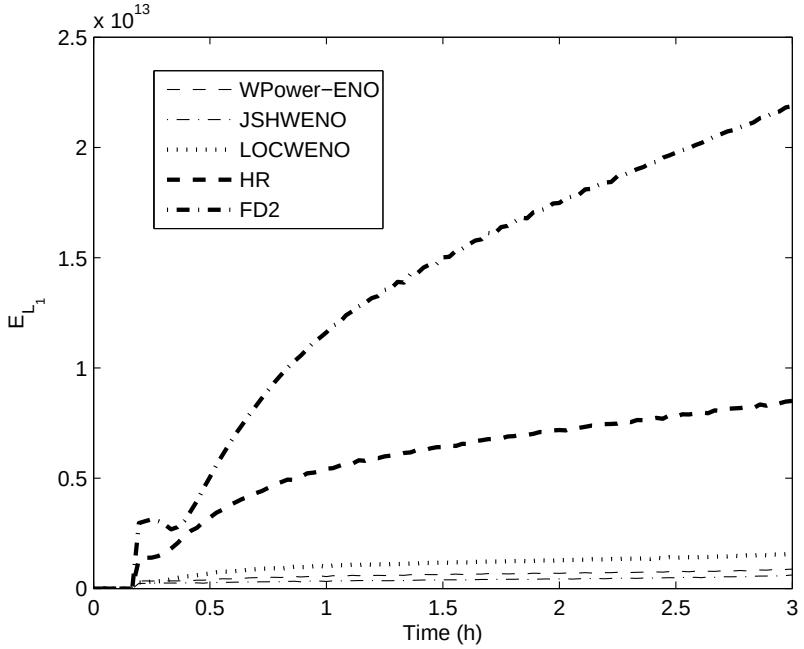


Figure 4.7: Evolution of the error L_1 norm with time for the various numerical methods ($\Delta L = 0.6\ \mu\mathrm{m}$).

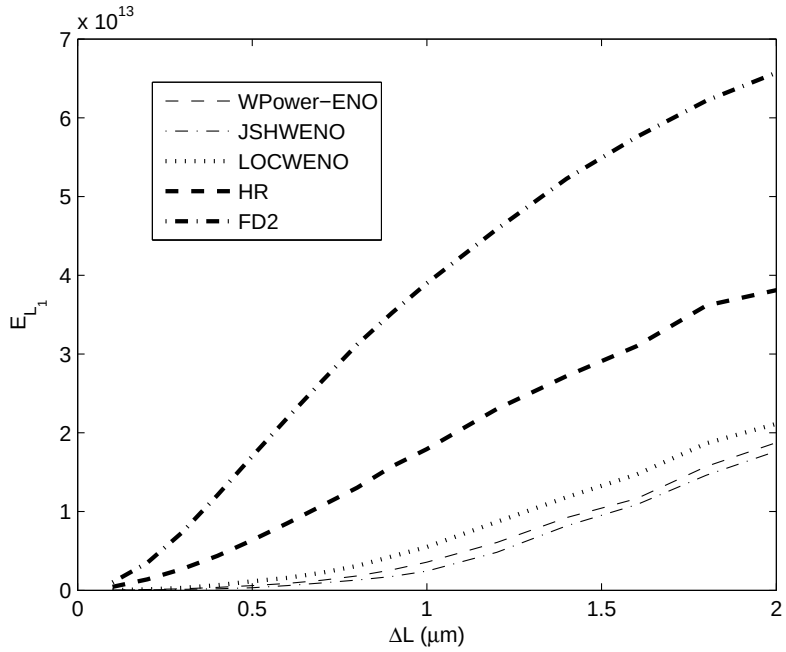


Figure 4.8: Error L_1 norm at the end of the batch versus ΔL for the various numerical methods.

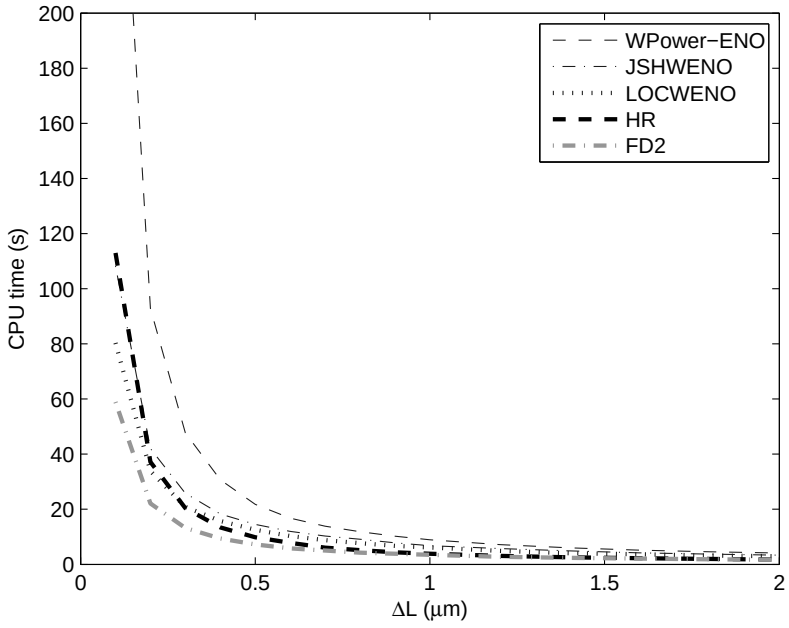


Figure 4.9: CPU time versus ΔL for the various numerical methods.

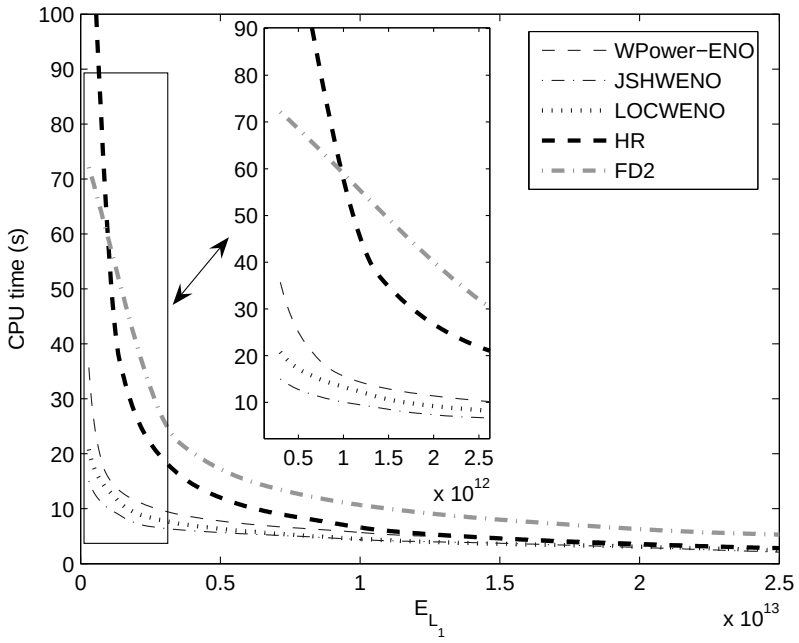


Figure 4.10: CPU time required for the various numerical methods for a given error L_1 norm at the end of the batch.

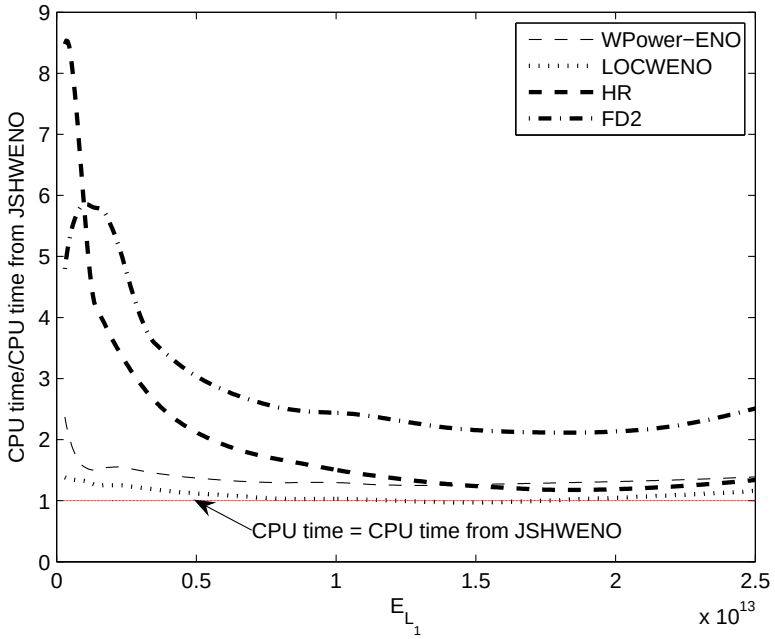


Figure 4.11: Relative CPU time for the various numerical methods with respect to CPU time from JSHWENO for a given error L_1 norm at the end of the batch.

4.4 Conclusions

This study proposed the use of WENO methods for the numerical solution of population balance models for crystallization processes. The LOCWENO, JSHWENO, and WPower-ENO methods were compared to standard discretization methods. In simulations of the polymorphic crystallization of L-glutamic acid, the WENO methods produced much less numerical diffusion and dispersion, with the LOCWENO and JSHWENO methods having the highest overall efficiency (that is, lowest CPU time for the same level of numerical accuracy). The L_1 self-convergence order which characterizes integral consistency and convergence was the highest for the JSHWENO method, followed by the other two WENO methods. These results recommend WENO methods for the simulation of crystallization processes, especially when the distributions are sharp and very high accuracy is desired. These methods combine very high order of accuracy with good convergence properties even in the presence of sharp variation in the size distributions.

Chapter 5

Temperature and Concentration Control Strategies

5.1 Introduction

The efficiency of downstream operations such as filtering and drying, and the efficacy of products can be dependent on the final crystal form and its size distribution. Hence, control of crystallization is an integral aspect in pharmaceutical industries. However, pharmaceutical crystallization process can be challenging to control due to variations in solution thermodynamics and kinetics induced by contaminants, complex nonlinear dynamics associated with non-ideal mixing and dendritic growth, and unexpected polymorphic phase transformation [129].

Most crystallization in pharmaceutical industries are designed and controlled based on trial-and-error experimentation, which can be time consuming and ex-

pensive. Recently in-process sensors have enabled the development of systematic first-principles (model-based) and direct design (measurement-based) approaches for the control of industrial crystallization processes [42].

The first-principles approach to crystallization control is the most widely studied [12, 128], where a model constructed from material and energy balances are used to optimize some function (e.g., mean crystal size, yield) of the crystal size distribution. Since small model uncertainties (e.g., kinetic parameters, solubility curves) can have a large effect on the crystal size and shape distribution of the product crystals, this approach requires the model to be sufficiently accurate and/or an appropriate measure to counteract the effect of the uncertainties.

The direct design approach, on the other hand, does not require first-principle models. Instead, this approach uses feedback control to follow a predetermined concentration-temperature curve in the metastable zone [41]. This approach requires in-situ concentration measurement in addition to temperature measurement. The concentration-temperature trajectory is suboptimal since it does not optimize any performance objective. Instead, this approach provides a constant tradeoff between the need to avoid excessive nucleation and to avoid overly long batch times (keeping the growth rate large) [42].

Until now, many studies have been done on non-polymorphic crystallization control which focused on controlling crystal size distribution and some characteristic functions derived from it. Recently, there has been a rapid growth in the study of polymorphism, with the desired objective being to produce one polymorph while

avoiding others. This is crucial in drug manufacturing, since different polymorphs of the same drug compound may have very different characteristics and may cause undesired side effect.

In this study, several control strategies for the polymorphic transformation of L-Glutamic acid from the metastable α-form to the stable β-form using first-principles approach are investigated and developed in Chapters 5 to 7. This chapter discusses the temperature control (T-control) strategy, of which objective is to obtain an optimal temperature profile as a function of time which maximizes a particular objective function, and concentration control (C-control) strategy, which implements optimal concentration-temperature trajectory. In the next section, the description of product quality, process constraints, and the parameter perturbations are given. Then the T-control and C-control strategies are developed, and followed by simulation studies which compare the performance and robustness of the two control strategies. This is followed by conclusions.

5.2 Product quality, process constraints, and parameter perturbations

For all control strategies studied in Chapters 5 to 7, two objectives are considered for the polymorphic transformation of α- to β-form crystals. The first objective is to maximize the mass of β-form crystals, which is equivalent to maximizing the

third-order moment or the yield of β-form crystals:

$$P_1 = (\mu_{\beta,3})_{t=t_f} \ . \tag{5.1}$$

The second objective is to minimize the ratio of the nucleated crystal mass to the seed crystal mass of β-form crystals, which can be written as

$$P_2 = \left(\frac{\mu_{\beta,3}^{\text{nucl}}}{\mu_{\beta,3}^{\text{seed}}} \right)_{t=t_f} , \tag{5.2}$$

where t_f is the batch time. The control problem is subject to the following inequality constraints:

$$T_{\text{min}} \leq T\left(t\right) \leq T_{\text{max}} \, , \tag{5.3}$$

$$C_{\text{sat},\beta}\left(t\right) \leq C\left(t\right) < C_{\text{sat},\alpha}\left(t\right) \ , \tag{5.4}$$

$$C\left(t_f\right) \leq C_{\text{max}}\left(t_f\right) \ , \tag{5.5}$$

where $T_{\text{min}} = 25°\text{C}$, $T_{\text{max}} = 50°\text{C}$ are the minimum and maximum temperatures due to the limitation of water bath heating/cooling. The inequality constraint (5.4) aims to avoid the nucleation and growth rate of α-form crystals and the dissolution of β-form crystals during the polymorphic transformation process. Finally, the final inequality constraint (5.5) ensures that the minimum yield required by economic considerations is satisfied.

In the polymorphic transformation process, both α- and β-form crystals are

seeded according to Gaussian distribution with parameter values given in Table 5.1 and the process is subject to two cases of parameter perturbations given in Table 5.2

Table 5.1: The parameters describing the seed distributions.

Seed	Mass $[\mathrm{g}]$	Mean crystals size $[\mu\mathrm{m}]$	Standard deviation of crystals size $[\mu\mathrm{m}]$
α	10.0	100.0	10.0
β	1.0	100.0	10.0

Table 5.2: Variations in model parameters for robustness study: Case 1 is the nominal model, Case 2 has slow nucleation and fast growth rate parameters for β-form crystals, and Case 3 has fast nucleation and slow growth rate parameters for β-form crystals.

Cases	$\ln(k_{b\beta,1})$	$\ln(k_{b\beta,2})$	$\ln(k_{g\beta,0})$	$\ln(k_{g\beta,2})$	g_β	$\ln(E_{g\beta})$
1	15.801	20.000	52.002	-0.251	1.047	12.078
2	15.758	19.961	53.200	-0.280	1.100	12.060
3	15.842	20.036	50.883	-0.240	1.019	12.070

5.3 T-control and C-control strategies

The most widely studied approach for the optimal control of non-polymorphic crystallization processes has utilized T-control in which the temperature trajectory has been computed from the optimization of an objective function based on an offline model with nominal parameters [128]. This is the most commonly used method found in literatures and has been implemented in pharmaceutical industry to maximize crystal size and minimize the coefficient of variation [161]. In this study, the design of T-control strategy comprises the minimization of the following objective

function:

$$J_{\mathrm{T-control}} = \min_{\mathbf{U}} \left(P - P_d\right)^2 , \tag{5.6}$$

where P and P_d are the predicted and desired final product quality (which corresponds to either Eq. (5.1) or (5.2)), respectively, $\mathbf{U} = \left[u_0, \ldots, u_{N-1}\right]^T$, u_k is the input value (i.e. the crystallizer temperature) at the kth sampling instance, and N is the total samples in a batch. The above minimization problem is subject to process model and inequality constraints (5.3) to (5.5). To implement the T-control strategy, the temperature-time trajectory is parameterized as a first-order spline with 18 time intervals and differential evolution (DE) [84, 151] technique is utilized to solve the above minimization problem.

In many experimental and simulation studies of non-polymorphic batch crystallizations, the C-control strategy (Figure 5.1) has resulted in low sensitivity of the product quality to most practical disturbances and variations in kinetic parameters [41, 42, 50, 97, 133, 175]. Recently, the C-control strategy has been applied to polymorphic crystallizations to produce large crystals of any selected polymorph [80] and to ensure maximum productivity in polymorphic transformation process [64]. C-control can be interpreted as nonlinear state feedback control [63, 175], in which the nonlinear master controller acts on the concentration C as a measured state [159] to produce the setpoint temperature T_{set} as its manipulated variable. The difference between the calculated T_{set} and the measured temperature T is used by

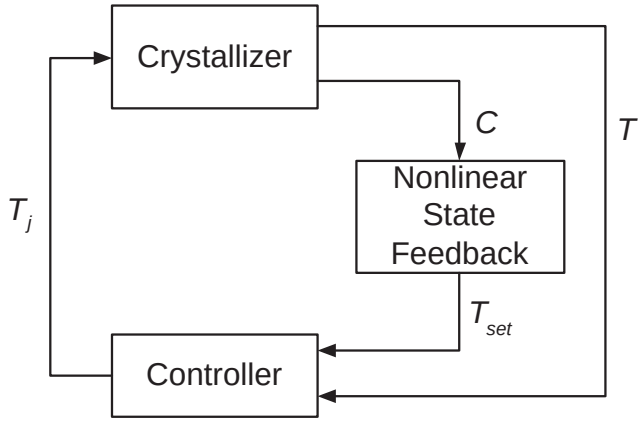


Figure 5.1: Implementation of C-control for a batch cooling crystallizer [175].

the slave controller to manipulate the jacket temperature T_j so that the deviation between T_{set} and T is reduced. Because the slave controller is just temperature control of a mixed tank, and the batch dynamics are relatively slow, any reasonably tuned proportional-integral controller will result in accurate tracking of T_{set}.

Hence, the main idea of C-control is to obtain a suitable parametrization for concentration-temperature trajectory and use it to calculate temperature setpoint throughout the crystallization run. In this study, the concentration-temperature trajectory is obtained by applying the optimal temperature-time trajectory from T-control to the nominal model. Then, a set of equations is used to parametrize the concentration-temperature trajectory, and is utilized to calculate temperature setpoint during crystallization process. The parameterization of the concentration-temperature trajectory corresponding to the product quality (5.1) is as follows (Figure 5.2):

(1) Initialize the temperature as $T_0 = 50°C$.

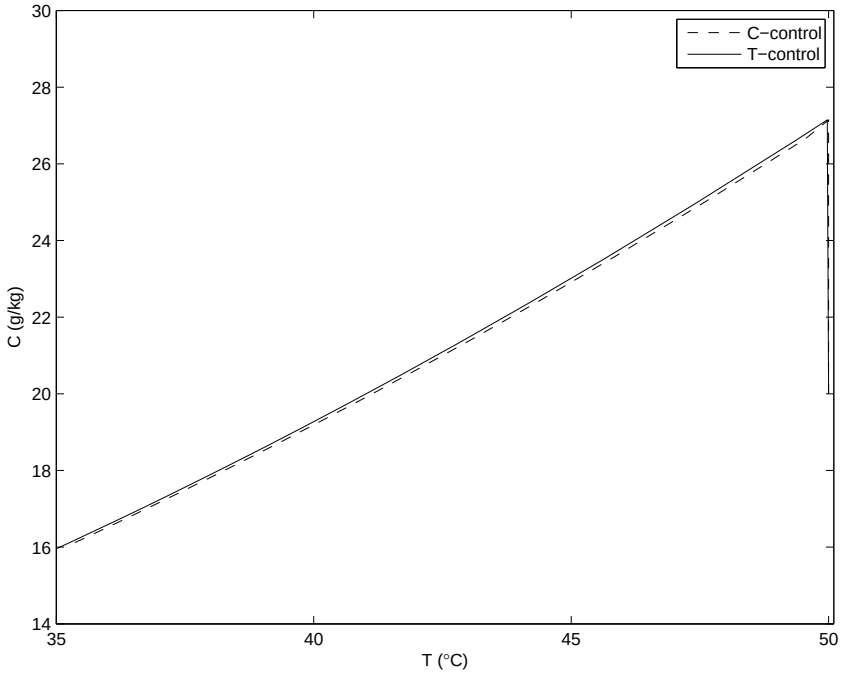


Figure 5.2: Concentration-temperature trajectory corresponding to product quality (5.1) obtained from T-control and C-control strategies.

(2) Calculate the temperature for the current sampling time T_k according to:

$$T_k = \min\left\{\max\left[1.003\, T_{\mathrm{ref},k}, 25\right], 50\right\}\,, \tag{5.7}$$

where

$$T_{\mathrm{ref},k} = \frac{-a_{\alpha,2} + \sqrt{a_{\alpha,2}^2 - 4a_{\alpha,1}\left(a_{\alpha,3} - C_k\right)}}{2a_{\alpha,2}}\,, \tag{5.8}$$

$a_{\alpha,i}$ are the parameters for the saturation concentration of α-form crystals (see Table 3.3), and C_k is the measurement of the solution concentration at the current sampling time.

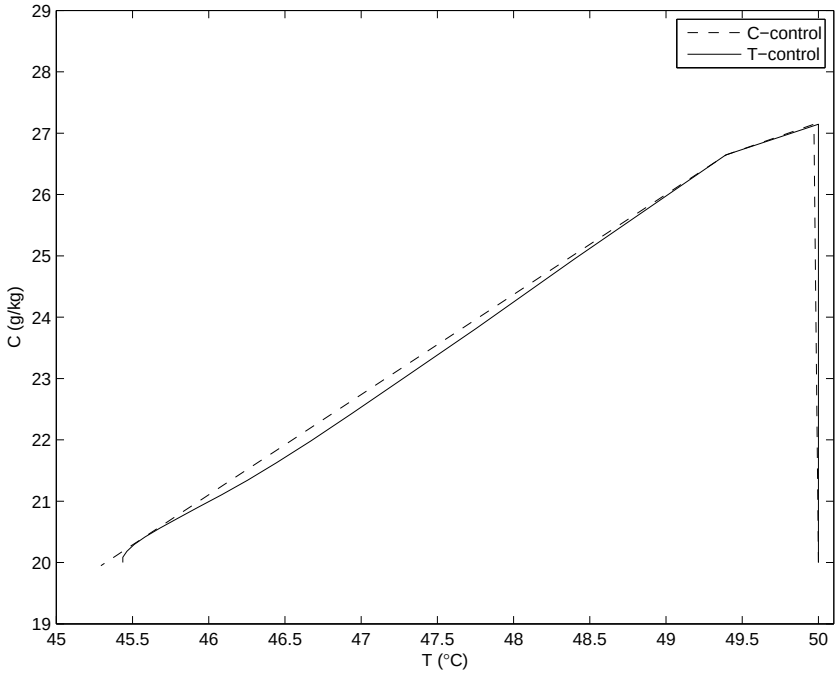


Figure 5.3: Concentration-temperature trajectory corresponding to product quality (5.2) obtained from T-control and C-control strategies.

On the other hand, the parameterization of the concentration-temperature trajectory corresponding to the product quality (5.2) is described below (Figure 5.3):

(1) Initialize the temperature as $T_0 = 50°C$ and set $mode = 1$, where $mode$ is a parameter used in the following step.

(2) Calculate $T_{\mathrm{ref},k}$ by Eq. (5.8) and obtain the temperature for the current sampling time T_k according to:

(a) If the α-form crystals are still present and

- If $T_{\mathrm{ref,k}} < 50$ and $mode = 1$, set $T_k = 50$.
- Otherwise, set $T_k = T_{\mathrm{ref,k}}$ and $mode = 2$.

(b) If the α-form crystals are absent, set

$$T_k = \min\left\{\max\left[T_{k-1} + 0.6125\left(C_k - C_{k-1}\right), 25\right], 50\right\}, \quad (5.9)$$

where T_{k-1} and C_{k-1} are the temperature and concentration measurements at the previous sampling time, respectively.

5.4 Simulation results

In the polymorphic transformation, both α- and β-form crystals are seeded according to a Gaussian distribution with parameter values given in Table 5.1. The initial solute concentration C_0 and maximum final solute concentration $C_{\max}(t_f)$ are 20 $\mathrm{g/kg}$ with a default batch time t_f is 3 hours which is extended if the inequality constraint (5.5) is not satisfied at that time. The sampling time is ten minutes. The optimization of two types of product qualities, P_1 in Eq. (5.1) and P_2 in Eq. (5.2), are considered, which from here onwards will be called objective J_1 and objective J_2, respectively. It is assumed that the process is subject to two cases of parameter perturbations given in Table 5.2.

The resulting temperature and concentration trajectories for T-control, C-control, and the optimal trajectories for objective J_1 are given in Figures 5.4 to 5.6. The optimal results are obtained using the T-control strategy by assuming that the parameter perturbations are known. The resulting values of P_1 for both control strategies and its optimal value are tabulated in Table 5.3. The optimal control trajectory for this

objective is very close to the solubility curve of α-form crystals (see Figure 5.4a) due to the slow growth rate of β-form crystals relative to the dissolution rate of α-form crystals. As a result, the optimal solution is to maximize the supersaturation with respect to the solubility of the β-form crystals while operating between the two solubility curves. When there is no plant-model mismatch (Figure 5.4), all control strategies produce similar results which are very close to the corresponding optimal one. As can be seen from Figure 5.5, the T-control strategy is not robust for the modelling error given by Case 2, because the resulting temperature trajectory deviates significantly from the optimal trajectory and the resulting P_1 value is 17% lower than the optimal one. Furthermore, for Case 3 (Figure 5.6), the T-control strategy violates one of the constraint most of the time during the batch. On the other hand, the C-control strategy shows a very good robustness in all cases where it produces P_1 values within 1% of the optimal ones. The robustness of the C-control strategy for this objective is in agreement with those reported by Hermanto et al. [64].

Objective J_2 is more sophisticated than objective J_1. For objective J_1, the purpose is to maximize the yield of β. Physically, this can be done when the nucleation and growth rates of β-form crystals are maximized. On the other hand, objective J_2 is equivalent to maximizing the yield of β-form crystals while trying to simultaneously minimize its nucleation. This results in maintaining a tradeoff between the nucleation and growth rates of β-form crystals. For objective J_2, the temperature and concentration trajectories obtained by all control strategies are shown in Figures 5.7 to 5.9 and the corresponding P_2 values are tabulated in Table 5.4. The opti-

mal solute concentration at the end of the batch is equal to the predefined $C_{\max}(t_f)$ for all three sets of parameters. For solute concentration and temperature considered here, the nucleated mass of β-form crystals always increases at a faster rate than the seed mass of β-form crystals. As the β-form crystals nucleate and grow, the ratio of nucleated crystal mass to seed crystal mass of β-form crystals always increases. As a result, any value of $C_{\max}(t_f)$ lower that the value specified by its upper bound constraint at 20 g/kg would increase the objective J_2 and would not be optimal. The performance of the T-control strategy for Case 2 is poor (Figure 5.8), which results in P_2 value 39% higher than the optimum value. In addition, implementing T-control strategy in Case 3 (Figure 5.9) needs the extension of batch time to 4.5 hours in order to satisfy the inequality constraint (5.5). For Cases 2 and 3, the P_2 values obtained by the C-control strategy is better than those obtained by the T-control strategy. However, they are achieved at the cost of long batch time, where it requires about 50 hours (Case 2) and 5.8 hours (Case 3) to satisfy the inequality constraint (5.5).

Table 5.3: Values of the control objective P_1 obtained for the three sets of model parameters in Table 5.2.

Cases	T-control	C-control	optimal
1	0.3119	0.3099	0.3119
2	0.3478	0.4187	0.4195
3	0.2569	0.2630	0.2667

Table 5.4: Values of the control objective P_2 obtained for the three sets of model parameters in Table 5.2.

Cases	T-control	C-control	optimal
1	0.0381	0.0385	0.0381
2	0.0064	$0.0050^{\ddagger}$	0.0046
3	$0.0683^{\ddagger}$	$0.0679^{\ddagger}$	0.0659

‡These values are obtained after the batch time is extended to satisfy constraint (5.5)

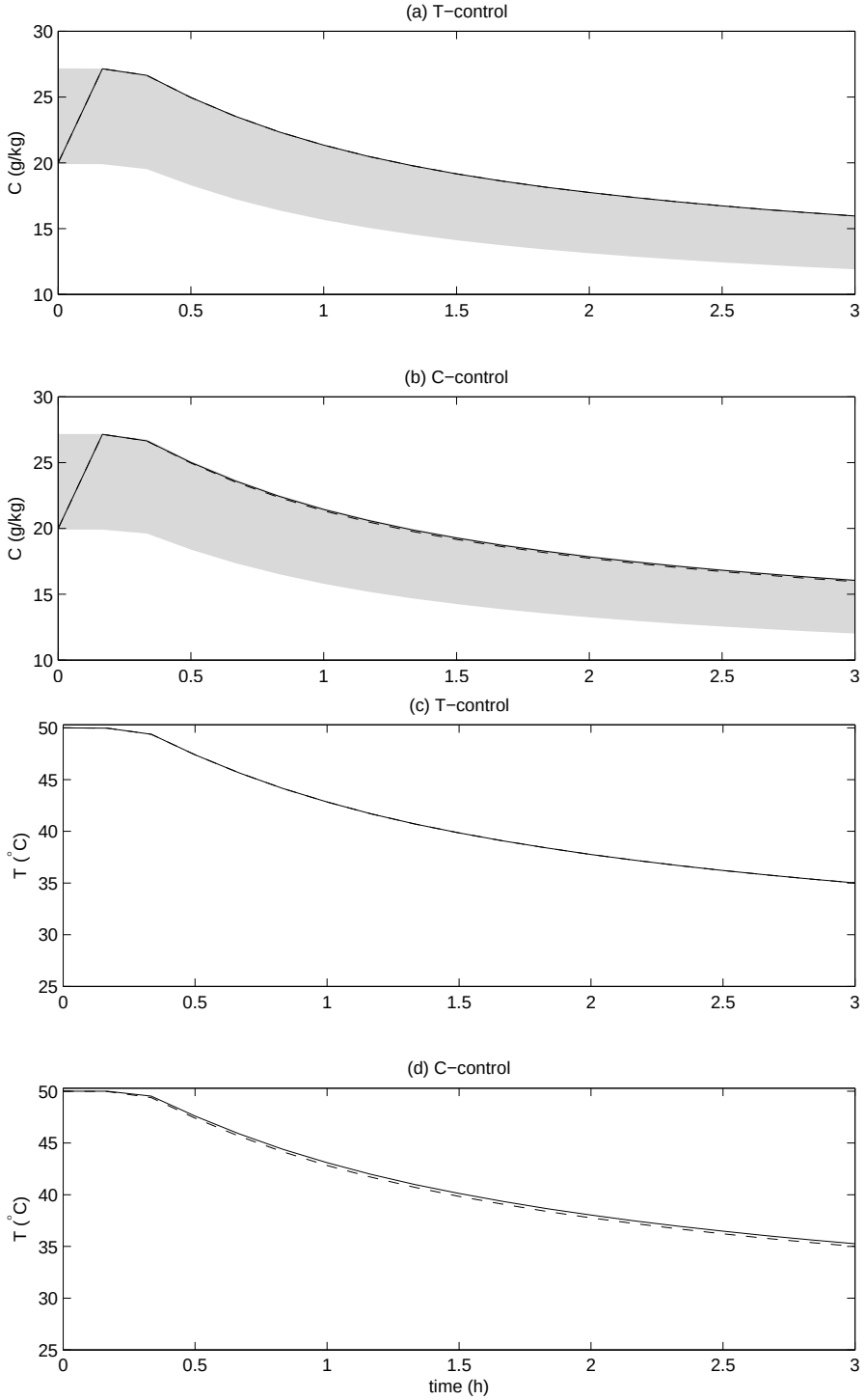


Figure 5.4: Concentration and temperature trajectories for Case 1 with objective J_1. The solid lines are trajectories corresponding to the two control strategies studied, the dashed lines are the optimal trajectories, and the shaded region indicates the inequality constraint (5.4) corresponding to the control strategies.

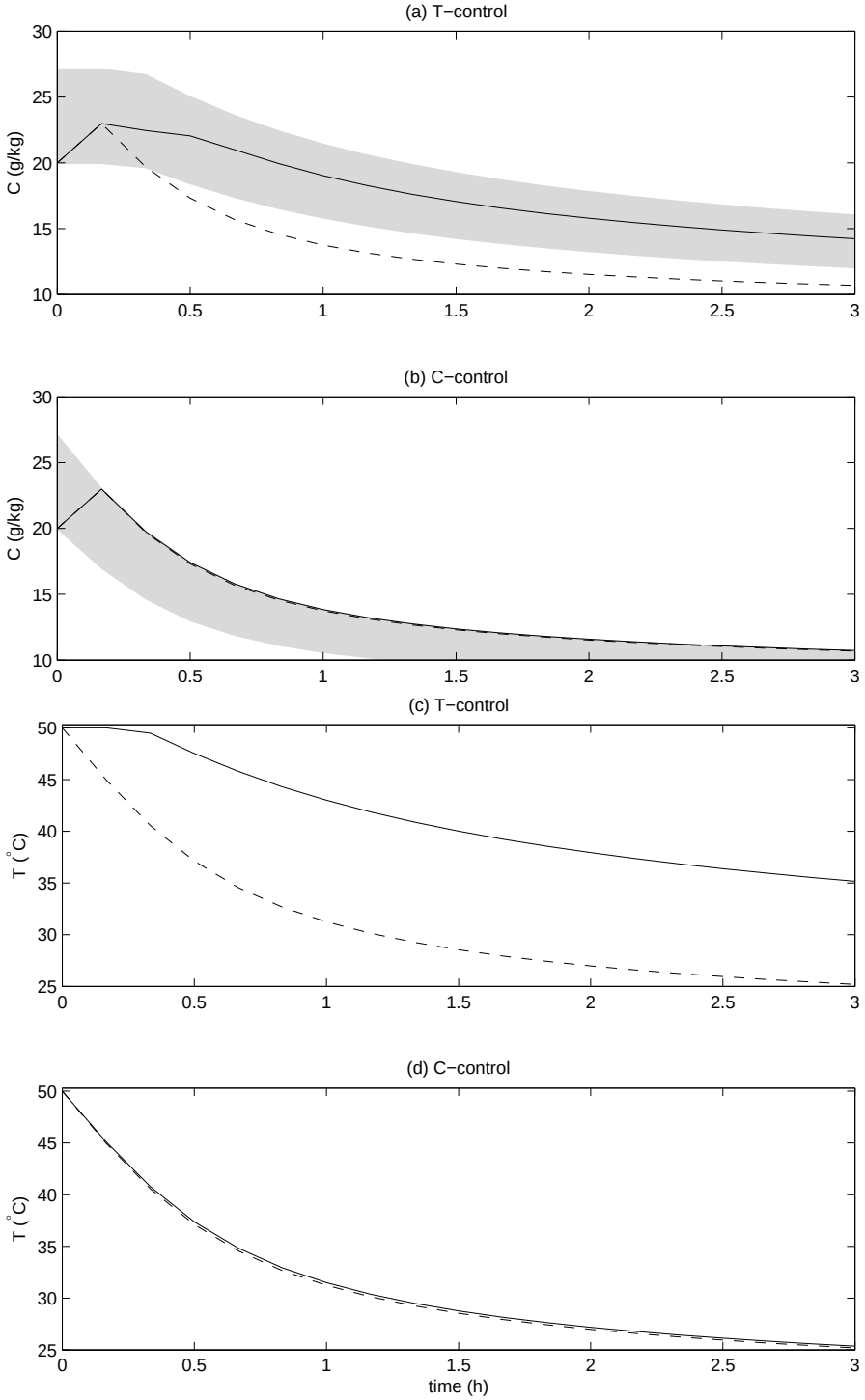


Figure 5.5: Concentration and temperature trajectories for Case 2 with objective J_1. The solid lines are trajectories corresponding to the two control strategies studied, the dashed lines are the optimal trajectories, and the shaded region indicates the inequality constraint (5.4) corresponding to the control strategies.

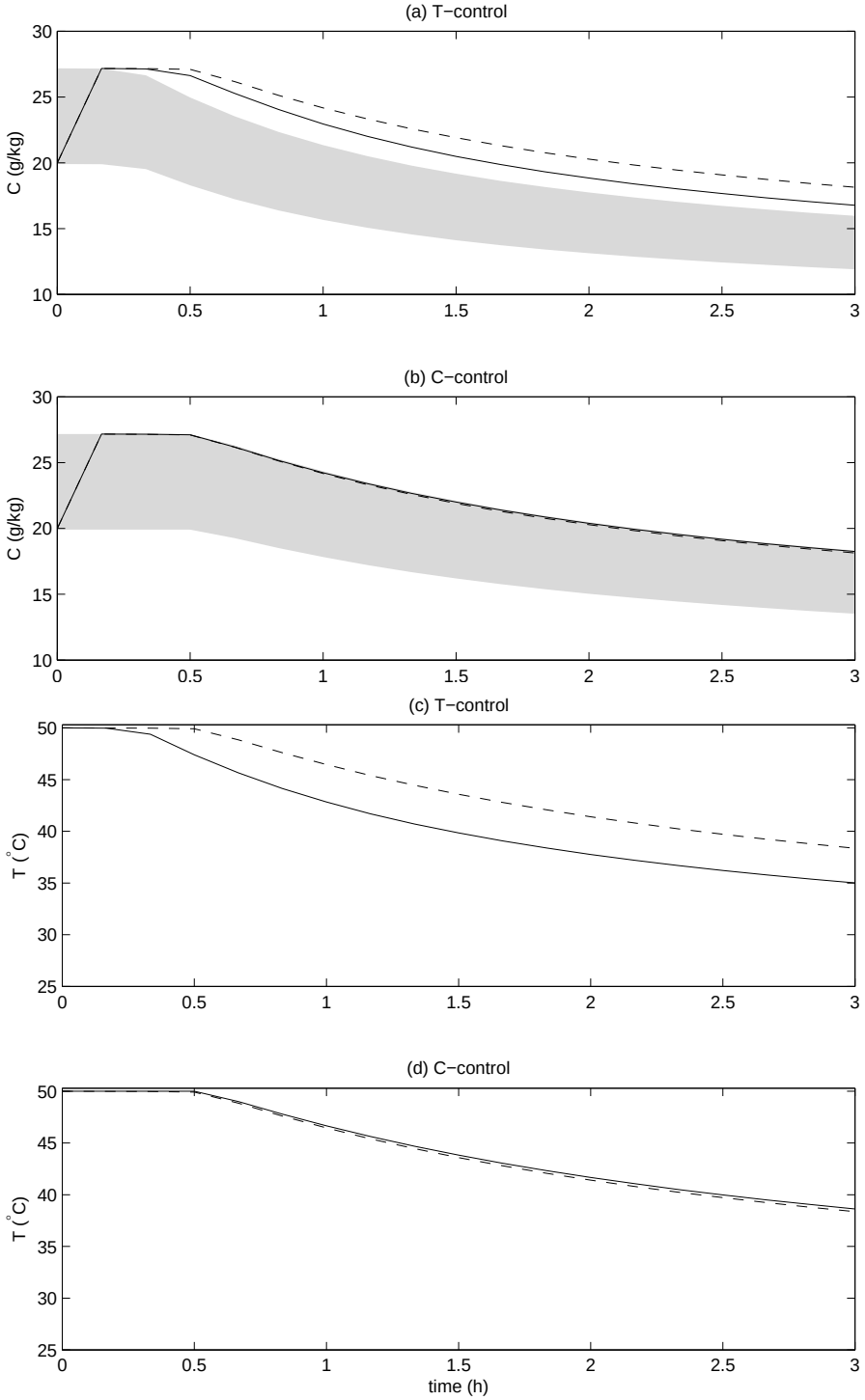


Figure 5.6: Concentration and temperature trajectories for Case 3 with objective J_1. The solid lines are trajectories corresponding to the two control strategies studied, the dashed lines are the optimal trajectories, and the shaded region indicates the inequality constraint (5.4) corresponding to the control strategies.

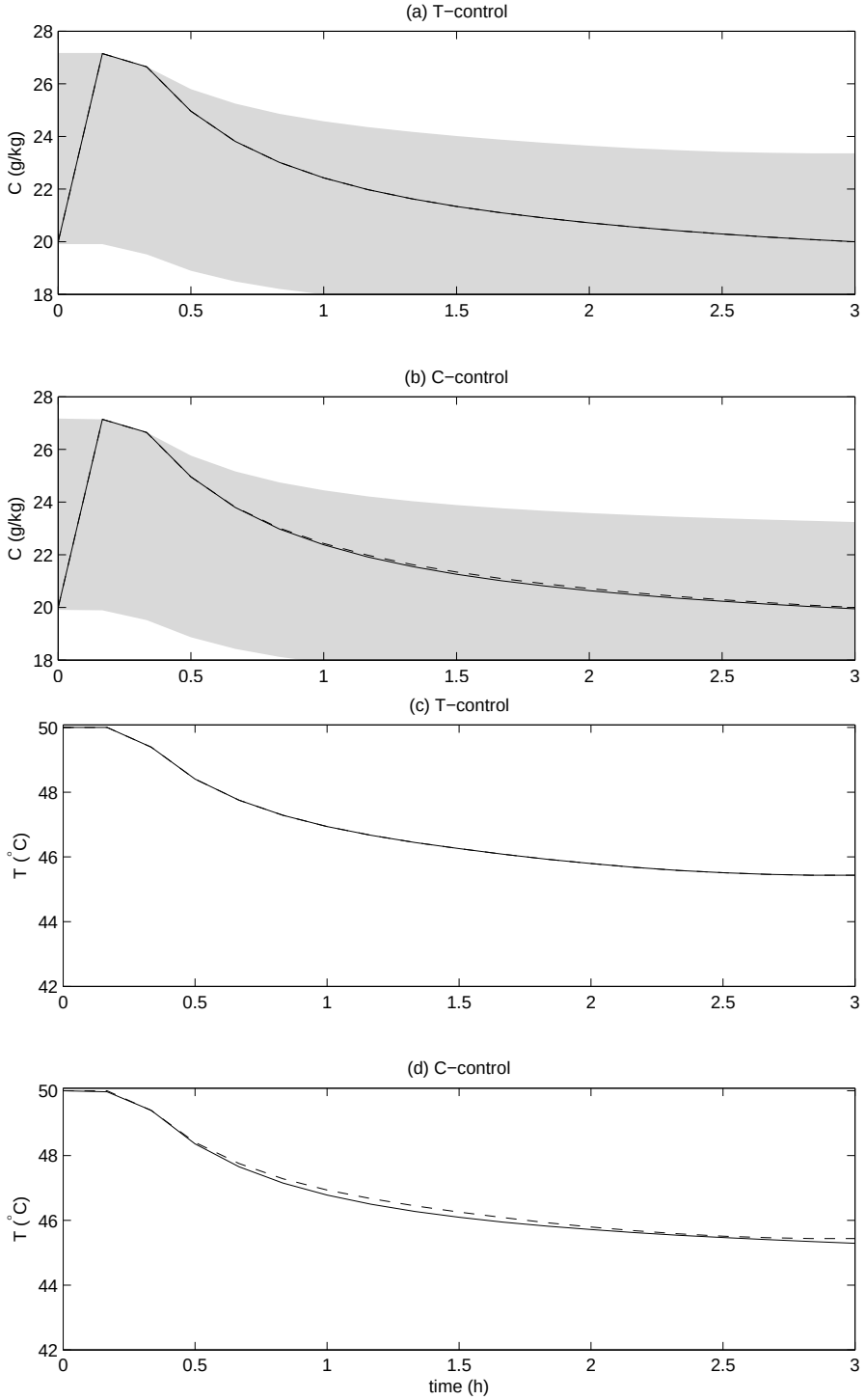


Figure 5.7: Concentration and temperature trajectories for Case 1 with objective J_2. The solid lines are trajectories corresponding to the two control strategies studied, the dashed lines are the optimal trajectories, and the shaded region indicates the inequality constraint (5.4) corresponding to the control strategies.

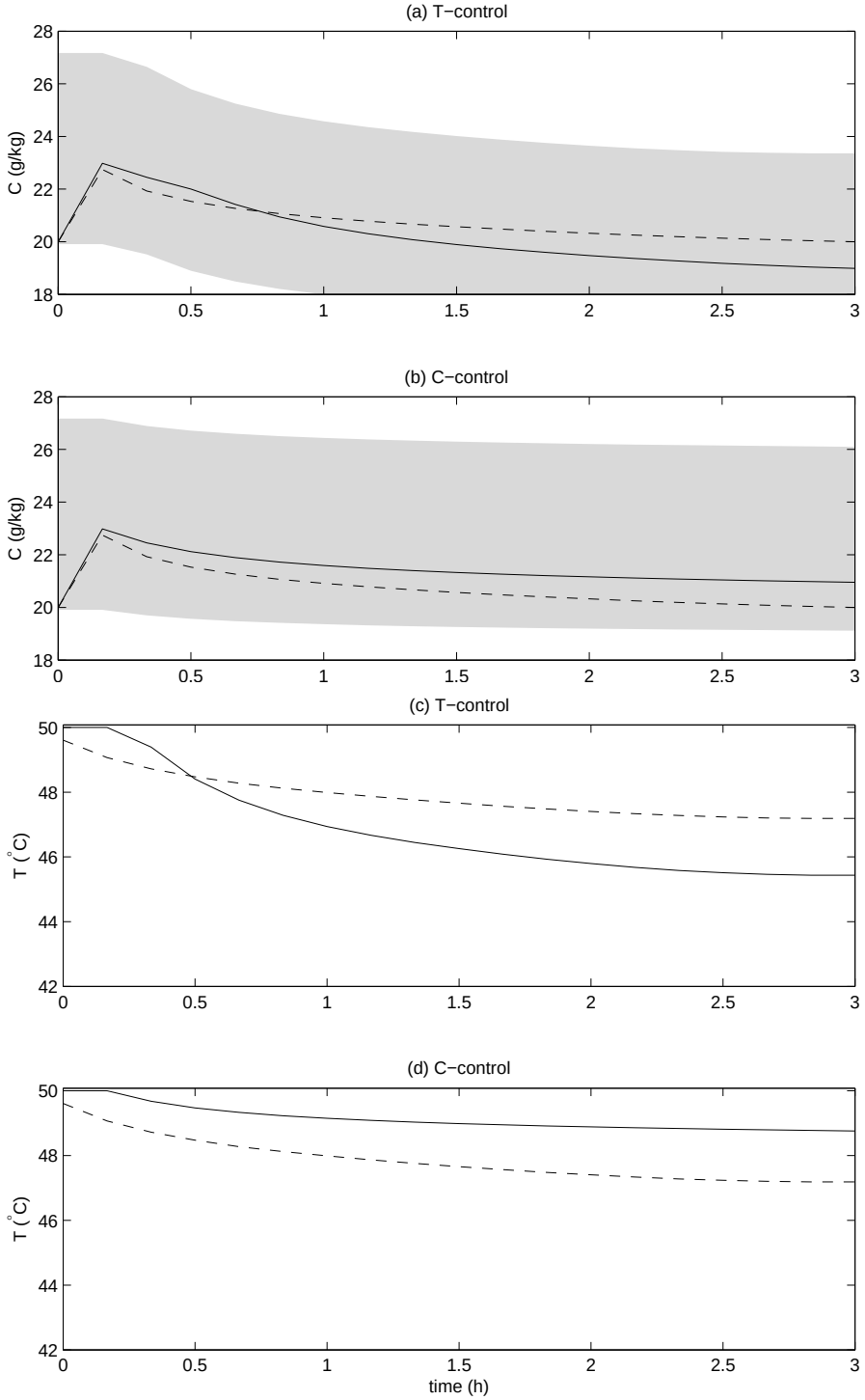


Figure 5.8: Concentration and temperature trajectories for Case 2 with objective J_2. The solid lines are trajectories corresponding to the two control strategies studied, the dashed lines are the optimal trajectories, and the shaded region indicates the inequality constraint (5.4) corresponding to the control strategies.

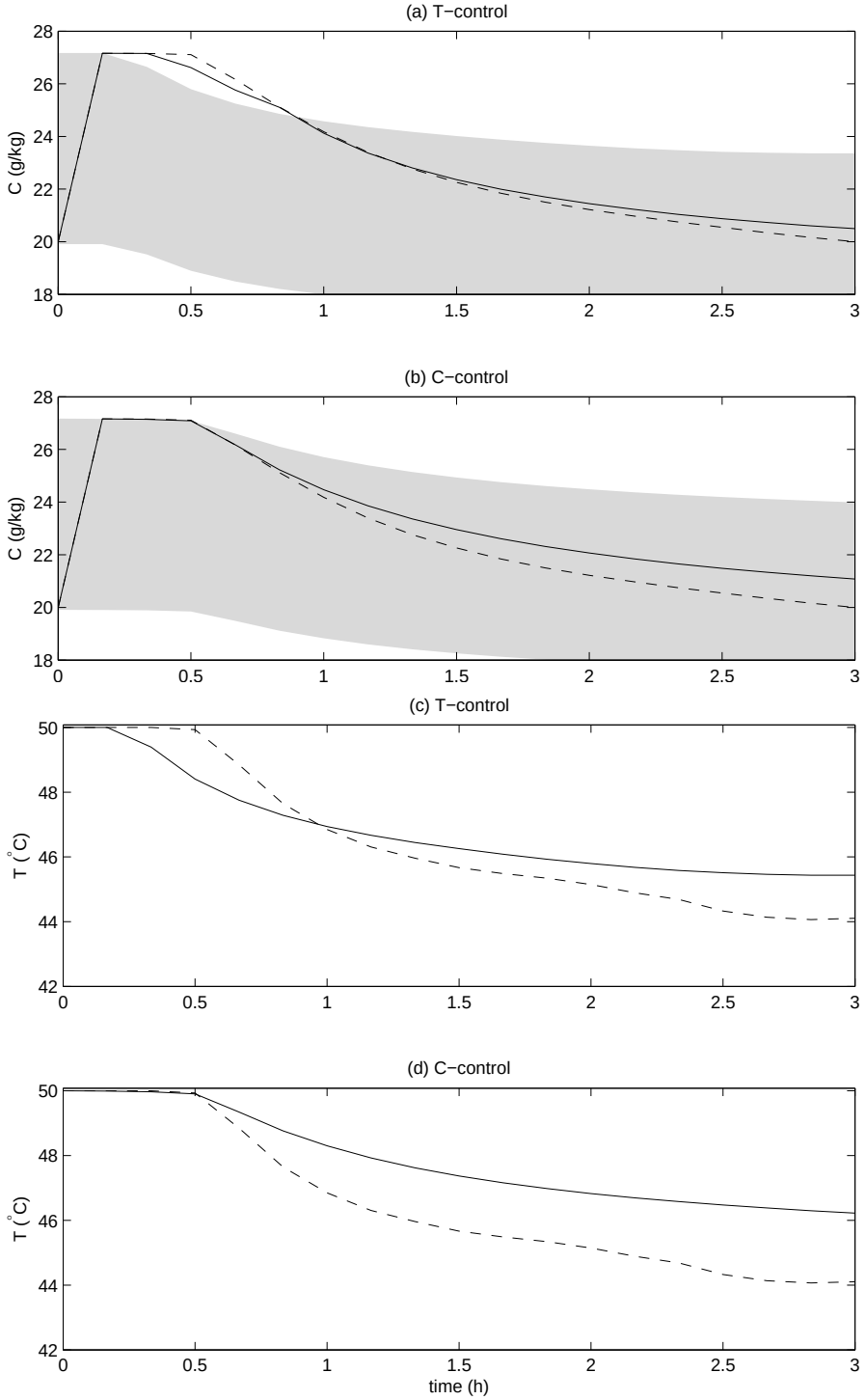


Figure 5.9: Concentration and temperature trajectories for Case 3 with objective J_2. The solid lines are trajectories corresponding to the two control strategies studied, the dashed lines are the optimal trajectories, and the shaded region indicates the inequality constraint (5.4) corresponding to the control strategies.

5.5 Conclusions

The T-control and C-control strategies, which are the popular control strategies in non-polymorphic crystallization processes, are discussed in this chapter. The robustness of these control strategies in the polymorphic transformation of L-glutamic acid from the metastable α-form to the stable β-form crystals is investigated. Two control objectives P_1 and P_2 are considered. The first objective aims to maximize the third moment of β-form crystals, which implies the yield of β-form crystals, whereas the second objective is to minimize the nucleated crystal mass to seed crystal mass ratio of β-form crystals.

From simulation results, it is shown that T-control is very sensitive to parameter perturbations, especially for Case 2, which results in 17% and 39% deviation from the optimal values of P_1 and P_2, respectively. On the other hand, C-control performs very robustly for objective J_1, which produces almost identical result to the optimal values for all cases. For objective J_2, the C-control strategy obtains better P_2 values than T-control, although longer batch times are required for Cases 2 and 3 to satisfy the yield constraint (5.5).

Chapter 6

Nonlinear Model Predictive Control Strategy

6.1 Introduction

Although T-control is the most widely adopted control strategy for crystallization process, it has become well-known that it can be very sensitive to variations in the kinetic parameters [13, 133], as also evidenced by the simulation results given in Chapter 5. While C-control can provide better robustness than the T-control, it is not capable of handling the input and output constraints commonly encountered during process operation. From the simulation results in Chapter 5, it is evident that C-control may require a very long batch time because of its failure in meeting the yield constraint within the specified batch time.

To address the shortcoming of T- and C-control strategies, the powerful model

predictive control (MPC) [19, 34, 44, 62, 110, 122, 127] techniques are considered in this chapter. Despite the high impact of MPC in academic research and industrial practice, its application to solution crystallization processes has been rather limited [35, 79, 113, 131, 155]. One contribution considered the effects of uncertainties on the closed-loop performance of nonlinear model predictive control (NMPC) applied to crystallization processes [113]. As in many other papers, the method of moments was utilized to simplify the population balance equations which are partial differential equations (PDEs) to a set of ordinary differential equations (ODEs) in terms of the moments. The NMPC optimization problem was solved using nonlinear programming and the states were estimated using an extended Kalman filter (EKF).

To the author's knowledge, there is no published result on the implementation of NMPC to a polymorphic crystallization, which is more challenging for a number of reasons. First, the phase equilibria and crystallization kinetics are more complicated. Second, the method of moments heavily used in past control algorithms for crystallization processes does not apply during a polymorphic transformation, so that the full PDEs need to be solved. As a consequence, the computation time required increases considerably which prohibits the straightforward application of nonlinear programming. In this study, a practical NMPC strategy based on extended predictive self-adaptive control (EPSAC) [32, 34, 70, 134, 156] is developed for the polymorphic transformation of L-glutamic acid from the metastable α-form to the stable β-form. To implement the proposed NMPC strategy, an unscented Kalman

filter (UKF) [74–78] is utilized to estimate the unmeasurable states. The performance and robustness of the proposed design is compared with T-control, C-control, and a standard NMPC algorithm in a numerical study.

This chapter is organized as follows. The next section describes the system representation and elaborates on the NMPC strategy based on EPSAC. This is followed by a brief description of the UKF which is utilized for the state estimation. Then the performance and robustness of the the proposed NMPC algorithm is compared to the T-control and C-control strategies. This is followed by conclusions.

6.2 System representation and NMPC strategy

The optimal control problem to be solved online at every sampling instance in the NMPC algorithm can be represented as

$$\min_{u_k} J(x_k, u_k) \tag{6.1}$$

subject to

$$x_k = f(x_{k-1}, u_{k-1}) + w_k \,, \tag{6.2}$$

$$d_k = d_{k-1} + \xi_k \,, \tag{6.3}$$

$$y_k = g(x_k, u_k) + d_k + v_k \,, \tag{6.4}$$

$$h(x_k, u_k) \leq 0 \,, \tag{6.5}$$

where J is the objective function; x_k, u_k, y_k, and d_k are the vectors of n_x system states, n_u inputs, n_y measured variables, and n_y unmeasured disturbances at the kth sampling instance; and w_k, ξ_k, and v_k are the vector of noises on the system states, unmeasured disturbances, and the measured variables. The system dynamics are described by the vector function f, the measurement equations by the vector function g, and the linear and nonlinear constraints for the system are described by the vector function h.

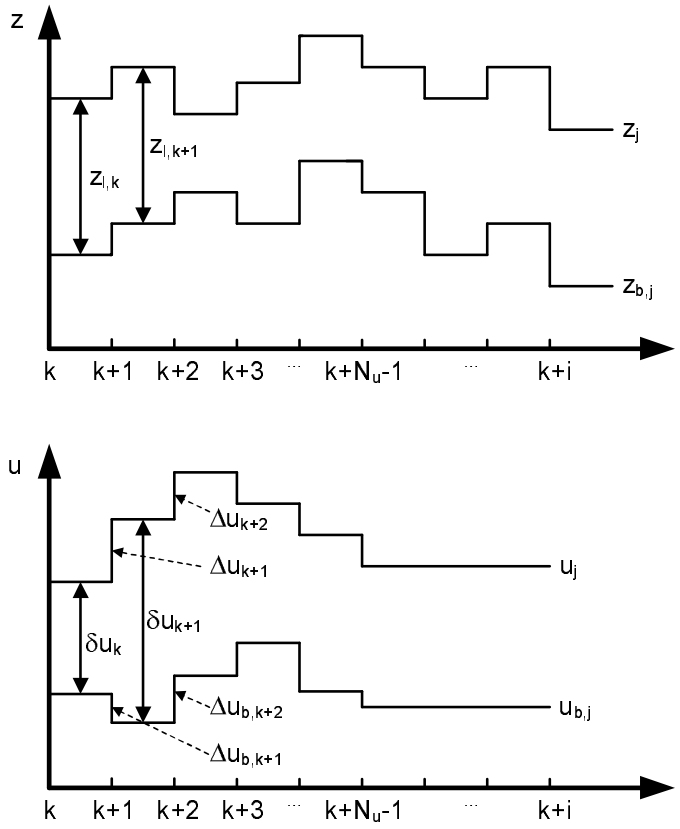


Figure 6.1: The variables decomposition in EPSAC.

The key idea of EPSAC is to approximate nonlinear process variables by iterative linearisation around future trajectories so that they converge to the same

nonlinear optimal solution [134]. For this purpose, the future sequence of the input variables u_{k+i} is considered as the sum of a predetermined future control scenario $u_{b,k+i}$ and the optimizing future control actions δu_{k+i} (Figure 6.1):

$$u_{k+i} = u_{b,k+i} + \delta u_{k+i}, \quad i = 0, 1, \ldots, N_u - 1, \tag{6.6}$$

where N_u is the control horizon and

$$\begin{aligned} \Delta u_{l,k} &= \delta u_k \,, \\ \delta u_{k+m} &= \sum_{j=0}^{m} \Delta u_{l,k+j} \,, \end{aligned}$$

Then the future trajectory of any process variables of interest (z_{k+i}) can be considered as being the cumulative result of two effects:

$$z_{k+i} = z_{b,k+i} + z_{l,k+i} \,, \tag{6.7}$$

where $z_{b,k+i}$ is calculated using the nonlinear model and predetermined sequence $u_{b,k+i}$. On the other hand, $z_{l,k+i}$ is obtained by:

$$\begin{aligned} z_{l,k+i} &= h_i \delta u_k + h_{i-1} \delta u_{k+1} + h_{i-2} \delta u_{k+2} + \cdots + h_{i-N_u+2} \delta u_{k+N_u-2} + \\ & \qquad h_{i-N_u+1} \delta u_{k+N_u-1} + \cdots + h_1 \delta u_{k+i-1} \,, \end{aligned} \tag{6.8}$$

where h_j is the jth impulse response coefficient. Noting that $\delta u_{k+N_u-1} = \delta u_{k+N_u} =$

$\cdots = \delta u_{k+i-1}$, the above equation can be written as

$$\begin{aligned} z_{l,k+i} &= h_i \delta u_k + h_{i-1}\delta u_{k+1} + h_{i-2}\delta u_{k+2} + \cdots + h_{i-N_u+2}\delta u_{k+N_u-2} + \\ &\quad (h_1 + h_2 + \cdots + h_{i-N_u+1})\,\delta u_{k+N_u-1} \\ &= h_i \delta u_k + h_{i-1}\delta u_{k+1} + h_{i-2}\delta u_{k+2} + \cdots + h_{i-N_u+2}\delta u_{k+N_u-2} + \\ &\quad g_{i-N_u+1}\delta u_{k+N_u-1}\,, \end{aligned} \tag{6.9}$$

where g_j is the jth step response coefficient.

For convenience, $z_{l,k+i}$ can be represented as a linear function of g_j and $\Delta u_{l,j}$:

$$\begin{aligned} z_{l,k+i} &= h_i \Delta u_{l,k} + h_{i-1}\sum_{j=0}^{1}\Delta u_{l,k+j} + h_{i-2}\sum_{j=0}^{2}\Delta u_{l,k+j} + \cdots + \\ &\quad h_{i-N_u+2}\sum_{j=0}^{N_u-2}\Delta u_{l,k+j} + g_{i-N_u+1}\sum_{j=0}^{N_u-1}\Delta u_{l,k+j} \\ &= (h_i + h_{i-1} + h_{i-2} + \cdots + h_{i-N_u+2} + g_{i-N_u+1})\,\Delta u_{l,k} + \\ &\quad (h_{i-1} + h_{i-2} + \cdots + h_{i-N_u+2} + g_{i-N_u+1})\,\Delta u_{l,k+1} + \cdots + \\ &\quad (h_{i-N_u+2} + g_{i-N_u+1})\,\Delta u_{l,k+N_u-2} + g_{i-N_u+1}\Delta u_{l,k+N_u-1} \\ &= g_i \Delta u_{l,k} + g_{i-1}\Delta u_{l,k+1} + \cdots + g_{i-N_u+1}\Delta u_{l,k+N_u-1}\,, \end{aligned} \tag{6.10}$$

Considering a batch process with the control horizon identical to the prediction horizon which covers from the next sampling time to the end of batch time denoted by $N_p = N_u = N - k$, where N is the total samples in a batch, the sequence of

$z_{l,k+i}$ is

$$\begin{aligned} z_{l,k+1} &= g_1 \Delta u_{l,k} \,, \\ z_{l,k+2} &= g_2 \Delta u_{l,k} + g_1 \Delta u_{l,k+1} \,, \\ &\vdots \\ z_{l,N} &= g_{N-k} \Delta u_{l,k} + g_{N-k-1} \Delta u_{l,k+1} + \cdots + g_1 \Delta u_{l,N-1} \,, \end{aligned}$$

or

$$\mathbf{Z}_l = \mathbf{G}_l \Delta \mathbf{U}_l \,, \tag{6.11}$$

where $\mathbf{Z}_l = [z_{l,k+1}, z_{l,k+2}, \cdots, z_{l,N}]^T$, $\Delta \mathbf{U}_l = [\Delta u_{l,k}, \Delta u_{l,k+1}, \cdots, \Delta u_{l,N-1}]^T$, and

$$\mathbf{G}_l = \begin{bmatrix} g_1 & 0 & \cdots & 0 \\ g_2 & g_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ g_{N-k} & g_{N-k-1} & \cdots & g_1 \end{bmatrix} .$$

In summary, the future process variables of interest in the prediction horizon can be conveniently represented in matrix form as

$$\mathbf{Z} = \mathbf{Z}_b + \mathbf{G}_l \Delta \mathbf{U}_l \,, \tag{6.12}$$

where $\mathbf{Z} = [z_{k+1}, z_{k+2}, \cdots, z_N]^T$ and $\mathbf{Z}_b = [z_{b,k+1}, z_{b,k+2}, \cdots, z_{b,N}]^T$.

Now consider an objective function J_{NMPC} with single input ($n_u = 1$):

$$J_{\mathrm{NMPC}} = \min_{\Delta \mathbf{U}} \quad [\mathbf{P} - \mathbf{P}_d]^T \mathbf{W}_p [\mathbf{P} - \mathbf{P}_d] + \Delta \mathbf{U}^T \mathbf{W}_u \Delta \mathbf{U} \,, \tag{6.13}$$

where $\mathbf{P}$, $\mathbf{P}_d$, and $\Delta \mathbf{U}$ are the vectors of the product quality, desired product quality, and the change in input variables, respectively, given by

$$\begin{aligned} \mathbf{P} &= [p_{k+1}, p_{k+2}, \cdots, p_N]^T \,, \\ \mathbf{P}_d &= [p_{d,k+1}, p_{d,k+2}, \cdots, p_{d,N}]^T \,, \\ \Delta \mathbf{U} &= [\Delta u_k, \Delta u_{k+1}, \cdots, \Delta u_{N-1}]^T \,, \end{aligned}$$

and $\mathbf{W}_p$ and $\mathbf{W}_u$ are the weight matrices for the product quality and the change in input variables, respectively. Then $\mathbf{P}$ and $\Delta \mathbf{U}$ can be decomposed into

$$\mathbf{P} = \mathbf{P}_b + \mathbf{G}_{pl} \Delta \mathbf{U}_l \,, \tag{6.14}$$

$$\Delta \mathbf{U} = \Delta \mathbf{U}_b + \Delta \mathbf{U}_l \,, \tag{6.15}$$

where $\mathbf{G}_{pl}$ is the step response coefficient matrix corresponding to the product quality variable, and $\mathbf{P}_b$ is the product quality calculated using the nonlinear model with predetermined future inputs $\mathbf{U}_b = [u_{b,k}, u_{b,k+1}, \cdots, u_{b,N-1}]^T$, and

$$\begin{aligned} \Delta \mathbf{U}_b &= [\Delta u_{b,k}, \Delta u_{b,k+1}, \cdots, \Delta u_{b,N-1}]^T \,, \\ \Delta \mathbf{U}_l &= [\Delta u_{l,k}, \Delta u_{l,k+1}, \cdots, \Delta u_{l,N-1}]^T \,. \end{aligned}$$

Hence, the minimization problem becomes:

$$\begin{aligned} J_{\mathrm{NMPC}} &= \min_{\Delta \mathbf{U}_l} \left[(\mathbf{P}_b - \mathbf{P}_d) + \mathbf{G}_{pl}\Delta \mathbf{U}_l\right]^T \mathbf{W}_p \left[(\mathbf{P}_b - \mathbf{P}_d) + \mathbf{G}_{pl}\Delta \mathbf{U}_l\right] + \\ &\qquad \left[\Delta \mathbf{U}_b + \Delta \mathbf{U}_l\right]^T \mathbf{W}_u \left[\Delta \mathbf{U}_b + \Delta \mathbf{U}_l\right] \\ &= \min_{\Delta \mathbf{U}_l} \Delta \mathbf{U}_l^T \mathbf{G}_{pl}^T \mathbf{W}_p \mathbf{G}_{pl} \Delta \mathbf{U}_l + 2\left(\mathbf{P}_b - \mathbf{P}_d\right)^T \mathbf{W}_p \mathbf{G}_{pl} \Delta \mathbf{U}_l + \\ &\qquad \Delta \mathbf{U}_l^T \mathbf{W}_u \Delta \mathbf{U}_l + 2\Delta \mathbf{U}_b^T \mathbf{W}_u \Delta \mathbf{U}_l \\ &= \min_{\Delta \mathbf{U}_l} \Delta \mathbf{U}_l^T \Gamma \Delta \mathbf{U}_l + \psi^T \Delta \mathbf{U}_l \,, \end{aligned} \tag{6.16}$$

where

$$\begin{aligned} \Gamma &= \mathbf{G}_{pl}^T \mathbf{W}_p \mathbf{G}_{pl} + \mathbf{W}_u \,, \\ \psi &= 2\left[\left(\mathbf{P}_b - \mathbf{P}_d\right)^T \mathbf{W}_p \mathbf{G}_{pl} + \Delta \mathbf{U}_b^T \mathbf{W}_u\right]^T . \end{aligned}$$

The minimization is subject to the constraints $h(x_j, u_j) \leq 0, \forall j \geq k$, where k is the current sampling instance. For notational convenience, $h(x_j, u_j)$ is denoted as h_j, which can be decomposed into the base and linear part $h_j = h_{b,j} + h_{l,j}$. Therefore, the matrix form of the constraints in the prediction horizon is

$$\mathbf{H}_b + \mathbf{G}_{hl} \Delta \mathbf{U}_l \leq 0 \,, \tag{6.17}$$

where $\mathbf{G}_{hl}$ is the step response coefficient matrix corresponding to the constraints function h_j and $\mathbf{H}_b = [h_{b,k}, h_{b,k+1}, \cdots, h_{b,N}]^T$.

From the author's experience, when the constraints are highly nonlinear, han-

dling the inequality constraints (6.17) directly will sometimes cause difficulty for the quadratic programming (QP) employed for the optimization to find a feasible solution. Therefore, the soft-constraint approach[140] is utilized which replaces the minimization problem with

$$\min_{\Delta \mathbf{U}_l, \epsilon} J_{\text{sc,NMPC}} \tag{6.18}$$

subject to

$$\mathbf{H}_b + \mathbf{G}_{hl} \Delta \mathbf{U}_l \leq \epsilon \,, \tag{6.19}$$

$$\epsilon \geq 0 \,, \tag{6.20}$$

where $J_{\text{sc,NMPC}} = J_{\text{NMPC}} + \epsilon^T \mathbf{W}_\epsilon \epsilon + \epsilon^T \mathbf{w}_\epsilon$, ϵ is a vector of slack variables, $\mathbf{W}_\epsilon$ is a diagonal matrix of positive weight, and $\mathbf{w}_\epsilon$ is a vector of positive elements. This modified minimization problem can be written as

$$\begin{aligned} J^*_{\text{sc,NMPC}} &= \min_{\Delta \mathbf{U}_l, \epsilon} \Delta \mathbf{U}_l^T \Gamma \Delta \mathbf{U}_l + \psi^T \Delta \mathbf{U}_l + \epsilon^T \mathbf{W}_\epsilon \epsilon + \epsilon^T \mathbf{w}_\epsilon \\ &= \min_{\Delta \mathbf{U}_l, \epsilon} \begin{bmatrix} \Delta \mathbf{U}_l^T & \epsilon^T \end{bmatrix} \begin{bmatrix} \Gamma & 0 \\ 0 & \mathbf{W}_\epsilon \end{bmatrix} \begin{bmatrix} \Delta \mathbf{U}_l \\ \epsilon \end{bmatrix} + \\ &\qquad \begin{bmatrix} \psi^T & \mathbf{w}_\epsilon^T \end{bmatrix} \begin{bmatrix} \Delta \mathbf{U}_l \\ \epsilon \end{bmatrix} \\ &= \min_{\Pi} \Pi^T \Lambda \Pi + \tau^T \Pi \,, \end{aligned} \tag{6.21}$$

subject to

$$\begin{bmatrix} \mathbf{H}_b \\ 0 \end{bmatrix} + \begin{bmatrix} \mathbf{G}_{hl} & -\mathbf{I} \\ 0 & -\mathbf{I} \end{bmatrix} \Pi \leq 0\,. \tag{6.22}$$

where $\Pi = [\Delta \mathbf{U}_l^T, \epsilon^T]^T$, $\Lambda = \begin{bmatrix} \Gamma & 0 \\ 0 & \mathbf{W}_\epsilon \end{bmatrix}$, and $\tau = [\psi^T, \mathbf{w}_\epsilon^T]^T$.

To summarize, the procedure for implementing the NMPC strategy based on EPSAC for each sampling instance k is:

(1) Obtain $\mathbf{U}_b$ by the following method:

- if $k = 0$ and $iter = 1$, $\mathbf{U}_b$ is chosen from the nominal operating point which was used in the previous batches;
- if $k > 0$ and $iter = 1$, $\mathbf{U}_b$ is set as the $\mathbf{U}_{\text{optimal}}$ obtained in the previous sampling instance;
- if $iter > 1$, the updated $\mathbf{U}_b$ from the previous iteration is used;

where $iter$ is the iteration count.

(2) Given the estimated current system states, obtain $\mathbf{P}_b$ and $\mathbf{H}_b$ by using $\mathbf{U}_b$ as the input to the nonlinear process model (6.1) to (6.5).

(3) Obtain the step response coefficient matrices $\mathbf{G}_{pl}$ and $\mathbf{G}_{hl}$ by introducing a step change in δu.

(4) Obtain $\Pi^* = [\Delta \mathbf{U}_l^*, \epsilon^*]^T$ from the solution to the minimization problem (6.21) and (6.22), then update the elements of $\mathbf{U}_b$ using

$$u_{b,k+j} = u_{b,k+j} + \sum_{i=0}^{j} \Delta u_{l,k+i} \, ,$$

where $j = 0, \ldots, N-1+k$.

(5) Calculate $err = \left\| \begin{bmatrix} \mathbf{G}_{pl} \\ \mathbf{G}_{hl} \end{bmatrix} \Delta \mathbf{U}_l^* \right\|$. If err is greater than a specified tolerance*, $iter = iter + 1$, and go back to Step 1. Otherwise, set $\mathbf{U}_{\text{optimal}} = \mathbf{U}_b$ and implement the first element of $\mathbf{U}_{\text{optimal}}$ to the process.

6.3 Unscented Kalman filter

In practice, not all states can be measured and those unmeasured states need to be estimated from available measurements. The most widely known state estimator for nonlinear systems is the extended Kalman filter (EKF). Although the EKF maintains the computationally efficient recursive update form of Kalman filter (KF), it has limitations. First, EKF relies on the linearization of the nonlinear system dynamics. Hence, if the system is highly nonlinear, then the state estimates can be poor. At worst, it may cause the state estimates to diverge. Secondly, linearization can be applied only if the Jacobian matrix exists. This means that EKF may not be applied to discontinuous systems. Finally, computing the Jacobian matrix can be poorly

*1×10^{-4} was used in this study.

numerically conditioned for some processes.

Julier et al. [78] proposed an approach for filtering nonlinear systems to address the aforementioned problems by using what is now known as the unscented transformation (UT) [74–77]. The UT works by constructing a set of points, referred as a sigma point, which are deterministically chosen to have the same known statistics (e.g., means and covariance) as a given state estimate. Then, a specified nonlinear transformation is applied to each sigma point, and the unscented estimate is obtained by computing the statistics of the transformed set. The incorporation of UT into the KF framework is called the unscented Kalman filter (UKF) [78]. The following describes briefly the implementation of UKF based on the spherical simplex unscented transformation [75].

Consider the n ($= n_x+n_y$) dimensional augmented system states $x_{a,k} = [x_k, d_k]^T$ and recast the system equations (6.2) to (6.4) as

$$x_{a,k} = f_a(x_{a,k-1}, u_{k-1}) + w_{a,k-1}\,, \tag{6.23}$$

$$y_k = g_a(x_{a,k}, u_k) + v_k\,, \tag{6.24}$$

where

$$\begin{aligned} f_a(x_{a,k-1}, u_{k-1}) &= \begin{bmatrix} f^T(x_{k-1}, u_{k-1}) & d_{k-1}^T \end{bmatrix}^T, \\ g_a(x_{a,k}, u_k) &= g(x_k, u_k) + d_k\,, \\ w_{a,k-1} &= \begin{bmatrix} w_k^T & \xi_k^T \end{bmatrix}^T. \end{aligned}$$

A summary of the UKF procedure is below:

(1) Calculate sigma points $\chi_{i,k-1}$ for $i = 0, 1, \ldots, n+1$ by the spherical simplex unscented transformation as follows:

(a) Obtain the initial weight sequence by specifying the weight for the first sigma point $\chi_{0,k-1}$, W_0^o, which is a scalar weight for the mean value of the augmented system states. The initial weights for the rest of the sigma points are obtained as

$$W_i^\circ = \frac{1 - W_0^\circ}{n+1} \text{ for } i = 1, \ldots, n+1\,. \tag{6.25}$$

(b) For $j = 2, 3, \ldots, n$, generate the vector sequence by using the following equation:

$$X_i^j = \begin{cases} \begin{bmatrix} X_0^{j-1} \\ 0 \end{bmatrix} & \text{for } i = 0\,, \\ \begin{bmatrix} X_i^{j-1} \\ -\frac{1}{\sqrt{j(j+1)W_1^\circ}} \end{bmatrix} & \text{for } i = 1, \ldots, j\,, \\ \begin{bmatrix} 0_{j-1} \\ \frac{j}{\sqrt{j(j+1)W_1^\circ}} \end{bmatrix} & \text{for } i = j+1\,, \end{cases} \tag{6.26}$$

where

$$X_0^1 = 0; \quad X_1^1 = -\frac{1}{\sqrt{2W_1^\circ}}; \quad X_2^1 = \frac{1}{\sqrt{2W_1^\circ}}\,. \tag{6.27}$$

(c) Compute the unscaled sigma points according to:

$$\chi_{i,k-1}^\circ = \hat{x}_{a,k-1} + A^T X_i^n\,, \tag{6.28}$$

where $\hat{x}_{a,k-1}$ and A^T are the mean of the augmented states and the Cholesky decomposition of the augmented states covariance matrix $P_{x_{a,k-1}}$ at previous sampling time, respectively. If $k = 1$, $\hat{x}_{a,0} = E\left[x_{a,0}\right]$ and $P_{x_{a,0}}$=$E\left[\left(x_{a,0} - \hat{x}_{a,0}\right)\left(x_{a,0} - \hat{x}_{a,0}\right)^T\right] = A^T A$. Otherwise, $\hat{x}_{a,k-1}$ and $P_{x_{a,k-1}}$ are defined in Step 7.

(d) Calculate the scaled sigma points and their associated weights by

$$\chi_{i,k-1} = \chi_{0,k-1}^\circ + \gamma\left(\chi_{i,k-1}^\circ - \chi_{0,k-1}^\circ\right)\,, \tag{6.29}$$

$$W_i = \begin{cases} \frac{\left(W_0^\circ + \gamma^2 - 1\right)}{\gamma} & \text{for } i = 0\,, \\ \frac{W_i^\circ}{\gamma^2} & \text{for } i \neq 0\,, \end{cases} \tag{6.30}$$

where $0 < \gamma \leq 1$ is a scaling factor indicating the distance of the sigma points to their mean values.

(2) Propagate the sigma points through the augmented system state equations:

$$\chi_{i,k|k-1} = f_a\left(\chi_{i,k-1}, u_{k-1}\right), \text{ for } i = 0, 1, \ldots, n+1. \tag{6.31}$$

(3) Compute the predicted mean and covariance matrix of the augmented system states from

$$\hat{x}_{a,k|k-1} = \sum_{i=0}^{n+1} W_i \chi_{i,k|k-1}\,, \tag{6.32}$$

$$\begin{aligned} P_{x_{a,k|k-1}} = & \sum_{i=0}^{n+1} \left[W_i(\chi_{i,k|k-1} - \hat{x}_{a,k|k-1})\left(\chi_{i,k|k-1} - \hat{x}_{a,k|k-1}\right)^T\right] + \\ & (1-\gamma^2)\left(\chi_{0,k|k-1} - \hat{x}_{a,k|k-1}\right)\left(\chi_{0,k|k-1} - \hat{x}_{a,k|k-1}\right)^T + Q\,, \end{aligned} \tag{6.33}$$

where Q is the process noise covariance matrix.

(4) Propagate $\chi_{i,k|k-1}$ through the observation equation:

$$y_{i,k|k-1} = g_a(\chi_{i,k|k-1}, u_k). \tag{6.34}$$

(5) Compute the predicted mean and the covariance matrix of the measured vari-

ables:

$$\hat{y}_{k|k-1} = \sum_{i=0}^{n+1} W_i y_{i,k|k-1}\,, \tag{6.35}$$

$$\begin{aligned} P_{y_{k|k-1}} = & \sum_{i=0}^{n+1} \left[W_i \left(y_{i,k|k-1} - \hat{y}_{k|k-1}\right)\left(y_{i,k|k-1} - \hat{y}_{k|k-1}\right)^T\right] + \\ & \left(1-\gamma^2\right)\left(y_{0,k|k-1} - \hat{y}_{k|k-1}\right)\left(y_{0,k|k-1} - \hat{y}_{k|k-1}\right)^T + R\,, \end{aligned} \tag{6.36}$$

where R is the measurement noise covariance matrix.

(6) Compute the predicted cross-covariance matrix between the augmented system states and the measured variables from

$$P_{x_a y} = \sum_{i=0}^{n+1} W_i \left(\chi_{i,k|k-1} - \hat{x}_{a,k|k-1}\right)\left(y_{i,k|k-1} - \hat{y}_{a,k|k-1}\right)^T. \tag{6.37}$$

(7) Once the measurement y_k is available, correct the predictions according to Kalman filter equations:

$$\hat{x}_{a,k} = \hat{x}_{a,k|k-1} + K_k(y_k - \hat{y}_{k|k-1})\,, \tag{6.38}$$

$$P_{x_{a,k}} = P_{x_{a,k|k-1}} - K_k P_{y_{k|k-1}} K_k^T\,, \tag{6.39}$$

where the Kalman gain K_k is defined as $K_k = P_{x_a y} P_{y_{k|k-1}}^{-1}$.

6.4 Simulation results and discussion

In the polymorphic transformation, both α- and β-form crystals are seeded according to a Gaussian distribution with parameter values given in Table 5.1. The initial solute concentration C_0 and maximum final solute concentration $C_{\max}(t_f)$ are 20 $\mathrm{g/kg}$ with a default batch time t_f is 3 hours which is extended if the inequality constraint (5.5) is not satisfied at that time. The sampling time is ten minutes.

6.4.1 Description of specific control implementations

In order to implement the NMPC strategy, the partial differential equations (4.1) and (4.2) need to be discretized into a series of ODEs as in Eq. (6.2). For this purpose, JSHWENO discretization method detailed in Chapter 4 is employed. The resulting discretizations of $f_{\text{seed},i}$ and $f_{\text{nucl},i}$ with respect to L together with the solute concentration are considered as the system states. Of all the system states, only solute concentration is measured. Hence, the rest of the system states need to be estimated from available measurements. In this study, the following measurements are considered:

$$y = [\mu_{\alpha,1},\, \mu_{\alpha,2},\, \mu_{\beta,1},\, \mu_{\beta,2},\, X_\alpha,\, C,\, T]^T, \tag{6.40}$$

where X_α is the crystal concentration of α-form crystals. The first four variables (i.e., the first- and second-order moments of α- and β-form crystals) can be mea-

sured using the online high-speed imaging system developed by the pharmaceutical manufacturer GlaxoSmithKline [31]. The crystal concentration of α-form crystals can be measured by Raman Spectroscopy [115, 139]. Several online techniques are available for measuring the solution concentration such as conductivity or attenuated total reflection Fourier transform infrared spectroscopy (ATR-FTIR) [128, 160]. Temperature measurements are readily available using teflon-coated thermocouples.

The optimization of two different product qualities, $P_1 = (\mu_{\beta,3})_{t=t_f}$ in Eq. (5.1) and $P_2 = \left(\frac{\mu_{\beta,3}^{\text{nucl}}}{\mu_{\beta,3}^{\text{seed}}}\right)_{t=t_f}$ in Eq. (5.2), are considered, which from here onwards will be called objective J_1 and objective J_2, respectively. The tuning parameters for the NMPC strategy for both objectives are given in Table 6.1. The performance and robustness of the NMPC strategy to the perturbations in the kinetic parameters in Table 5.2 are compared with that of T-control, C-control, and quadratic dynamic matrix control with successive linearization (SL-QDMC). The T-control and C-control strategies are explained in the previous chapter, while the formulation of SL-QDMC is based on the quadratic dynamic matrix control (QDMC) by Garcia and Morshedi [43], with the successive linearization of the process model performed to obtain the dynamic matrix at every sampling instance. The constraints are handled in a similar way as in the NMPC strategy. Although the simulation results of the T- and C-control strategies have been discussed and provided in Chapter 5, the pertaining discussions are repeated here for the sake of easy reference.

6.4.2 Comparison results and discussion

For the first control objective J_1, the concentration and temperature trajectories for all four control strategies compared to the corresponding optimal trajectories, for the three sets of parameters, are shown in Figures 6.2 to 6.4.[†] The corresponding values of P_1 (which is proportional to the mass yield of β crystals) are in Table 6.2. When there is no plant-model mismatch (Figure 6.2), all control strategies except SL-QDMC produce similar results which are very close to the optimal solution. As discussed in Chapter 5, the T-control strategy is not robust for the modelling error given by Case 2, with the temperature trajectory deviating significantly from the optimal trajectory and the resulting P_1 value is 17% lower than the optimal one (Figure 6.3 and Table 6.2). For Case 3 (Figure 6.4), the T-control strategy violates one of the constraints most of the time during the batch. In contrast, the C-control strategy provides very good robustness in all cases, producing P_1 values within 1% of the optimal ones. The poor performance of the SL-QDMC for Cases 1 and 2 with P_1 values 13% and 31% lower than optimal may be accounted for by the high process nonlinearity and the closeness of the optimal solution to a constraint, the solubility curve of α-form crystals. This closeness to the constraint prevents the use of aggressive tuning parameters for SL-QDMC, otherwise the constraint is violated even for Case 1 with no model uncertainty. SL-QDMC results in good performance for Case 3 with a P_1 value within 2% of the optimal one, but violates a

[†]The optimal temperature trajectory was computed by applying T-control to the sets of parameters treated as known.

state constraint near the end of the batch (Figure 6.4c). The NMPC strategy shows good robustness for Case 3, but it is less robust for Case 2. Although the temperature and concentration trajectories for the NMPC strategy for Case 2 are different from the optimal trajectories (Figure 6.3), the P_1 value is nearly optimal (i.e., within 4%) for all three sets of parameters.

For objective J_2, the temperature and concentration trajectories obtained by the four control strategies are plotted in Figures 6.5 to 6.7 with P_2 values tabulated in Table 6.3. For Case 1 (Figure 6.5), all control strategies except SL-QDMC produce nearly the optimal P_2 value (Table 6.3). The performance of T-control for Case 2 is poor (Figure 6.6), with a P_2 value 39% higher than optimum. Implementing T-control in Case 3 (Figure 6.7) needs an extension of the batch time to 4.5 hours in order to satisfy the inequality constraint on the yield (5.5). The P_2 value obtained by C-control is much better than obtained by the T-control for Case 2 but only moderately better for Case 3 (Table 6.3). This improved robustness is achieved, however, by using a longer batch time, requiring about 50 hours (Case 2) and 5.8 hours (Case 3) to satisfy the inequality constraint (5.5). For Cases 2 and 3, the P_2 values of SL-QDMC are worse than C-control and NMPC but the difference is not nearly as big as in Case 1. For all three sets of parameters, SL-QDMC was able to satisfy all the constraints for the second objective within the specified batch time.

For the second objective, the NMPC strategy had the best performance and robustness among the four control strategies for both sets of perturbed parameters, with P_2 values within 7% from the optimal ones. While C-control and NMPC gave

nearly the same P_2 values and both satisfied all of the constraints during the entire batch, a clear advantage of NMPC was that it completed the batches within the specified batch time. Although the greater ease of implementation of C-control makes it easier to transfer to industry [97, 175], this simulation study demonstrates that there is room for improved robust performance and productivity by using a more sophisticated NMPC strategy.

Table 6.1: Tuning parameters for the NMPC strategy.

Values for objective J_1	Values for objective J_2
$\mathbf{W}_p = I$	$\mathbf{W}_p = I$
$(\mathbf{W}_u)_{i,i}{}^{\dagger} = 7\left[1 + 10\left(i-1\right)\right] \times 10^{-4}$	$(\mathbf{W}_u)_{i,i}{}^{\dagger} = \left[1 + 2\left(i-1\right)\right] \times 10^{-4}$
$\mathbf{W}_\epsilon = 10I$	$\mathbf{W}_\epsilon = 10I$
$\mathbf{w}_\epsilon = 10\left[1, 1, \ldots, 1\right]^T$	$\mathbf{w}_\epsilon = 10\left[1, 1, \ldots, 1\right]^T$
$W_0^\circ = 0.8$	$W_0^\circ = 0.8$
$\gamma = 0.1$	$\gamma = 0.1$

†The diagonal elements of matrices $\mathbf{W}_u$, where $i = 1, \ldots, N-k$

Table 6.2: Values of the control objective P_1 obtained for the three sets of model parameters in Table 5.2.

Cases	T-control	C-control	SL-QDMC	NMPC	optimal
1	0.3119	0.3099	0.2720	0.3117	0.3119
2	0.3478	0.4187	0.2881	0.4031	0.4195
3	0.2569	0.2630	0.2634	0.2666	0.2667

Table 6.3: Values of the control objective P_2 obtained for the three sets of model parameters in Table 5.2.

Cases	T-control	C-control	SL-QDMC	NMPC	optimal
1	0.0381	0.0385	0.0406	0.0384	0.0381
2	0.0064	0.0050‡	0.0053	0.0049	0.0046
3	0.0683‡	0.0679‡	0.0681	0.0679	0.0659

‡These values are obtained after the batch time is extended to satisfy constraint (5.5)

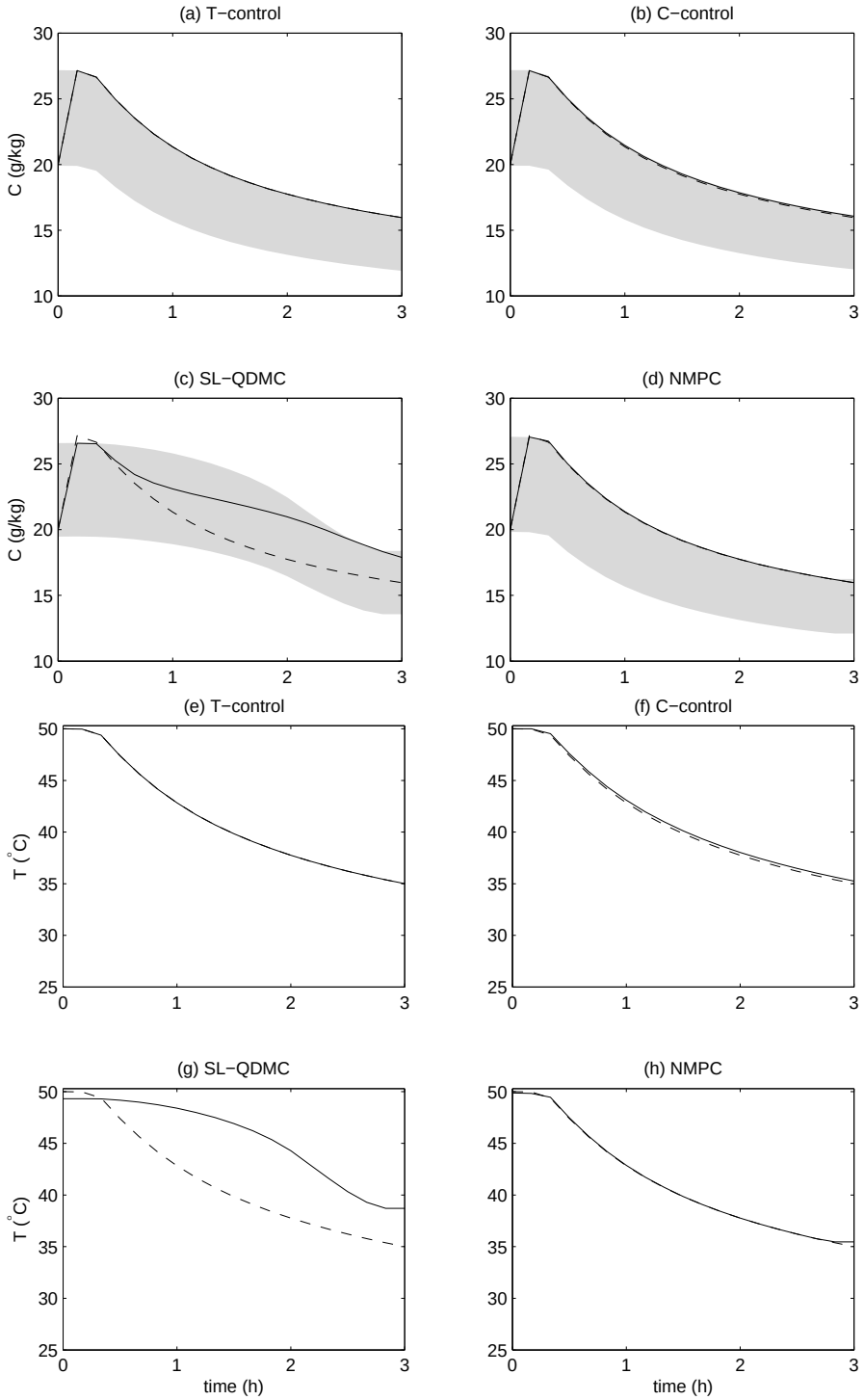


Figure 6.2: Concentration and temperature trajectories for Case 1 with objective J_1. The solid lines are trajectories corresponding to the four control strategies studied, the dashed lines are the optimal trajectories, and the shaded region indicates the inequality constraint (5.4) corresponding to the control strategies.

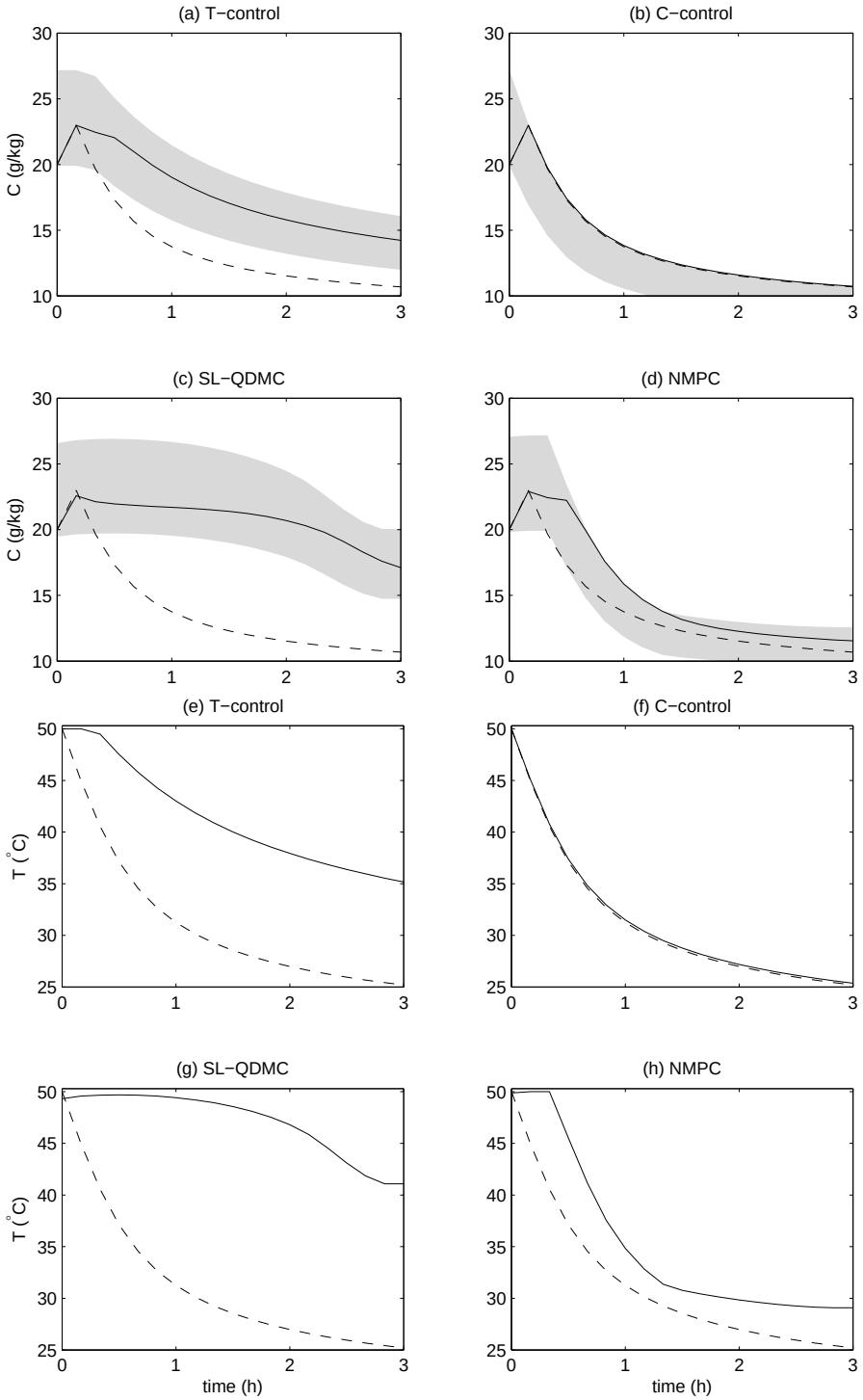


Figure 6.3: Concentration and temperature trajectories for Case 2 with objective J_1. The solid lines are trajectories corresponding to the four control strategies studied, the dashed lines are the optimal trajectories, and the shaded region indicates the inequality constraint (5.4) corresponding to the control strategies.

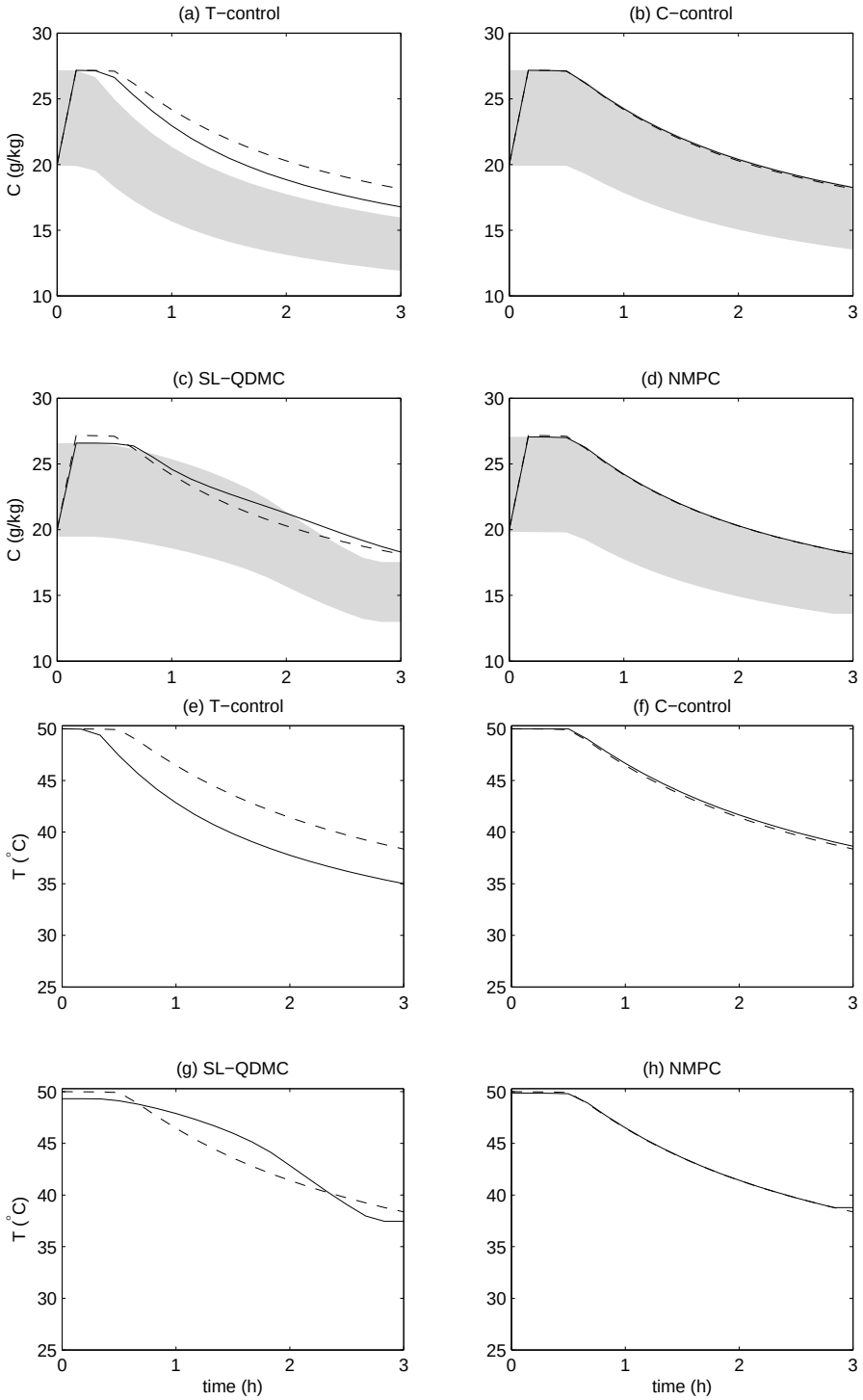


Figure 6.4: Concentration and temperature trajectories for Case 3 with objective J_1. The solid lines are trajectories corresponding to the four control strategies studied, the dashed lines are the optimal trajectories, and the shaded region indicates the inequality constraint (5.4) corresponding to the control strategies.

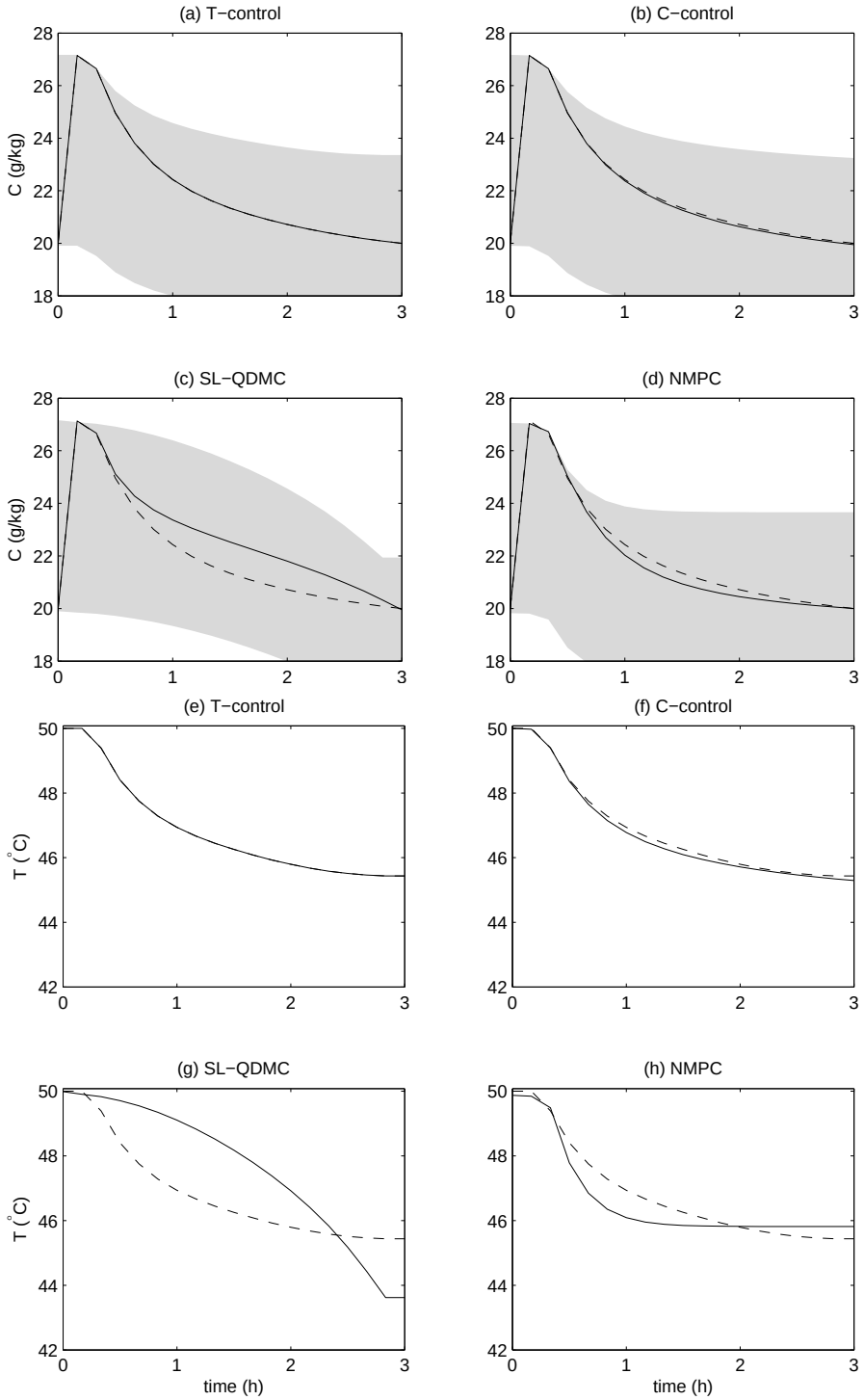


Figure 6.5: Concentration and temperature trajectories for Case 1 with objective J_2. The solid lines are trajectories corresponding to the four control strategies studied, the dashed lines are the optimal trajectories, and the shaded region indicates the inequality constraint (5.4) corresponding to the control strategies.

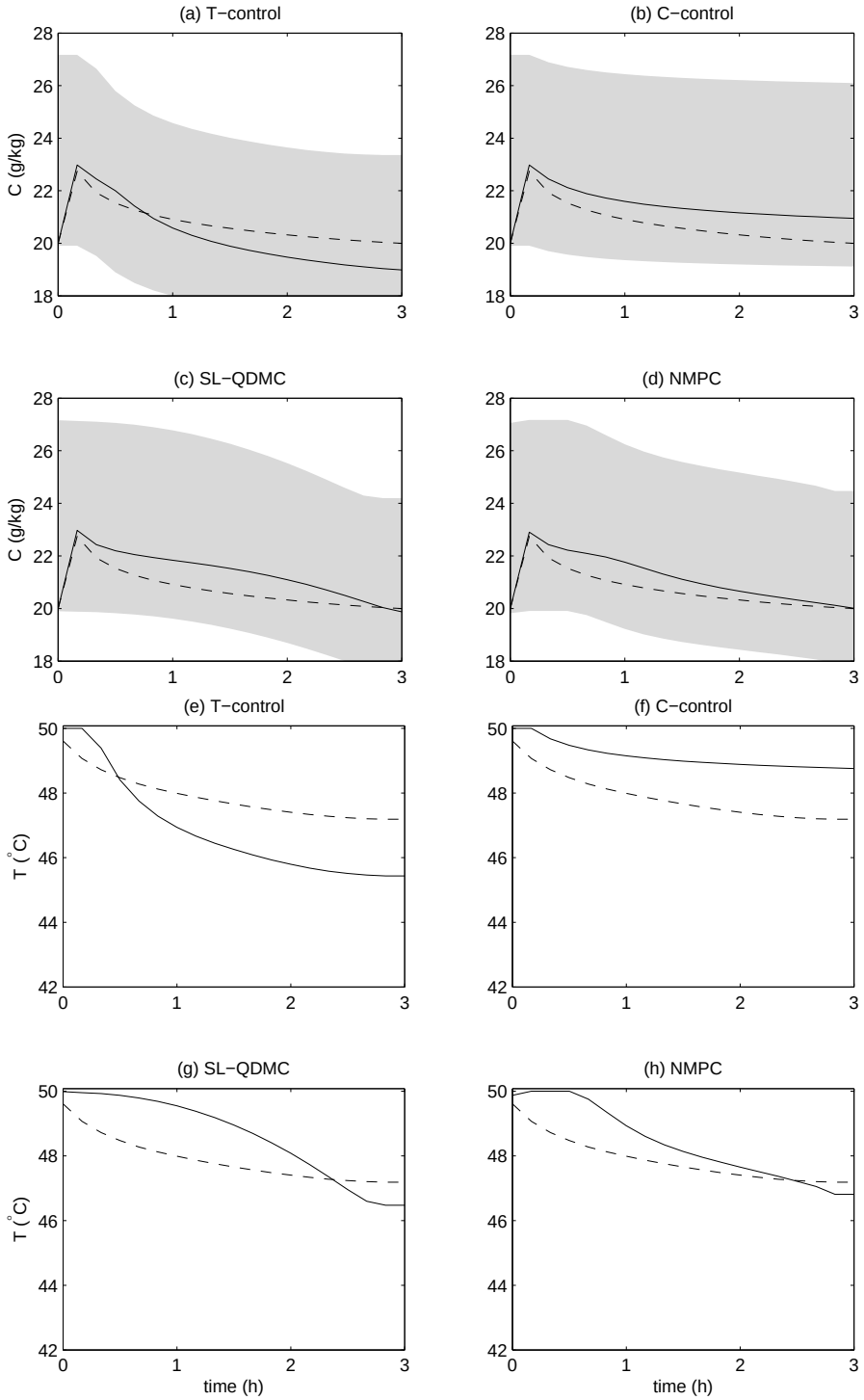


Figure 6.6: Concentration and temperature trajectories for Case 2 with objective J_2. The solid lines are trajectories corresponding to the four control strategies studied, the dashed lines are the optimal trajectories, and the shaded region indicates the inequality constraint (5.4) corresponding to the control strategies.

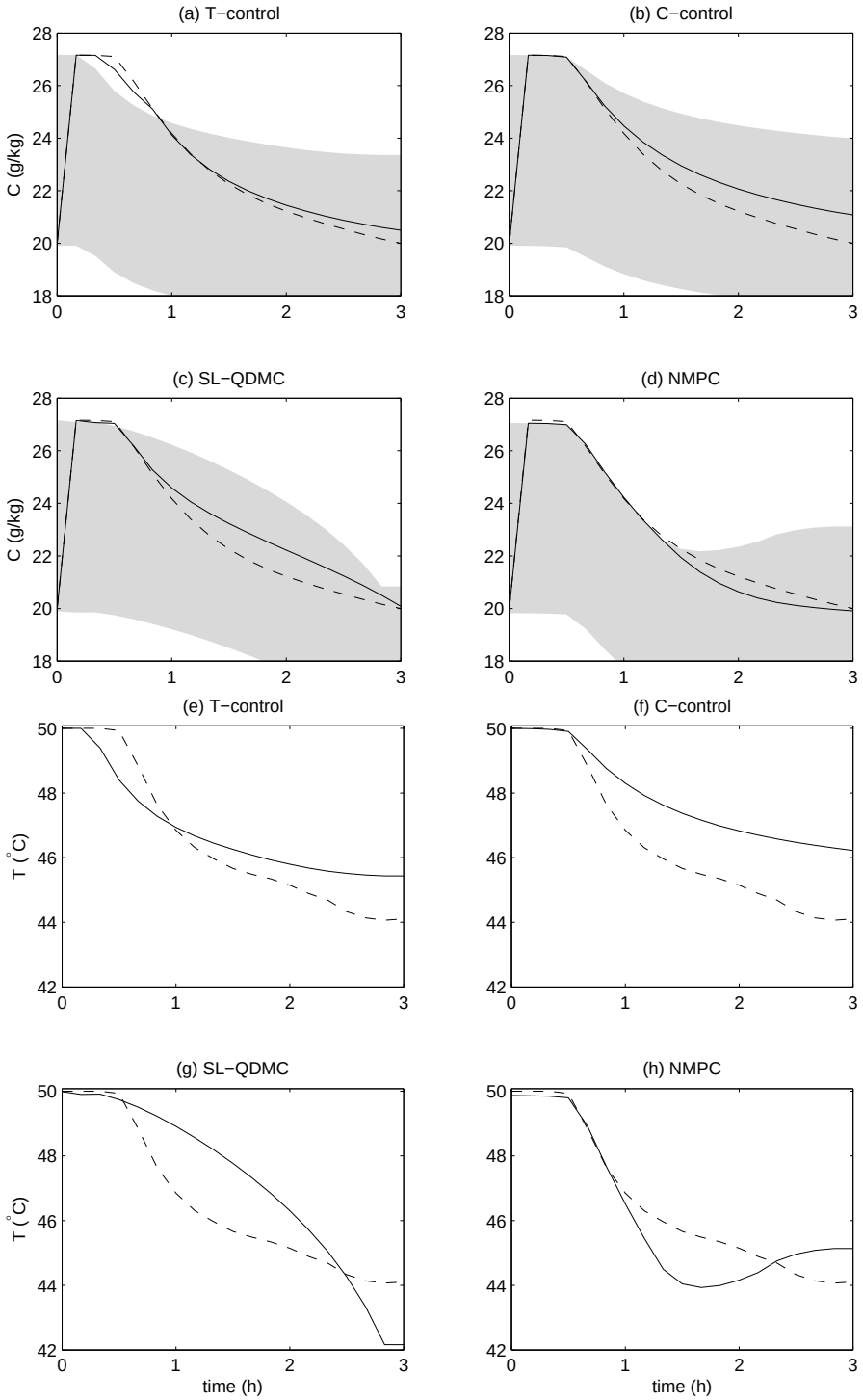


Figure 6.7: Concentration and temperature trajectories for Case 3 with objective J_2. The solid lines are trajectories corresponding to the four control strategies studied, the dashed lines are the optimal trajectories, and the shaded region indicates the inequality constraint (5.4) corresponding to the control strategies.

6.5 Conclusions

An NMPC strategy based on EPSAC and UKF was presented for batch polymorphic crystallization processes. A simulation study considered the control objectives of maximizing the yield of β-form crystals (P_1) and minimizing the ratio of nucleated crystal mass to seed crystal mass of β-form crystals (P_2). The performance and robustness of the NMPC strategy was compared to established control strategies, namely T-control, C-control, and SL-QDMC.

T-control was very sensitive to parameter perturbations, especially for Case 2, which results in 17% and 39% deviation from the optimal values of P_1 and P_2, respectively. C-control was robust for the maximization of yield, which produced almost identical results to the optimal values for three sets of parameters. While C-control satisfied all of the constraints and produced P_2 values better or similar to those of the other control strategies, the simulation study showed that C-control could take a very long batch time to satisfy the yield constraint (5.5). SL-QDMC performed very poorly, even when there is no plant-model mismatch, due to high process nonlinearity exacerbated by closeness of the optimal solution to a state constraint. The NMPC strategy showed good overall robustness for both objectives (within 4% and 7% of the optimal values, respectively) while satisfying all constraints within the specified batch time.

Chapter 7

Integrated Nonlinear MPC and Batch-to-Batch Control Strategy

7.1 Introduction

Due to the plant-model mismatch, optimal control obtained from offline process model is often suboptimal when applied to the real process. Exploiting the fact that batch processes are repetitive in nature, batch-to-batch control uses results from previous batches to iteratively compute the optimal operating conditions for each batch. Batch-to-batch control has been studied extensively in the past decade. Zafiriou et al. [173] proposed an approach for modifying the input sequence from batch-to-batch to deal with plant-model mismatch. Their approach is based on an analogy between the iteration during numerical optimization of an objective function and successive batches during the operation of the plant. Clarke-Pringle and MacGre-

gor [26] proposed a method to correct the manipulated variable trajectories from batch-to-batch with application to the optimization of molecular weight distribution in a polymerization process. The method uses errors between the measured and desired molecular weight distributions at the end of a batch to update the manipulated variable trajectories for the next batch. Lee and co-workers [89] presented the quadratic criterion-based iterative learning control (Q-ILC) approach for tracking control of batch processes based on a linear time-varying tracking error transition model. Doyle et al. [37] used batch-to-batch optimization to achieve the desired particle size distribution (PSD) target in an emulsion polymerization reactor. A simplified theoretical model is used as predictor, but the prediction is corrected using an updated PLS model that relates the manipulated variables to the error from the theoretical model prediction and the measured distribution. Xiong and Zhang [170] presented a recurrent neural network based ILC scheme for batch processes where the filtered recurrent neural network prediction errors from previous batches are added to the model predictions for the current batch and optimization is performed based on the updated predictions. Li et al. [94] presented a batch-to-batch optimal control based on recursively updated nonlinear model. In their approach, a batch-wise recursive nonlinear PLS algorithm was proposed to update the model after each batch.

With the ability of model predictive control (MPC) to respond to disturbances occuring during the batch and batch-to-batch control to correct any bias left uncorrected by the MPC, combining both methods to obtain better control performance

is possible. This integrated control strategy can combine information from the past error tracking signals with that from the current batch to adjust the manipulated variable trajectories more effectively in real time. If disturbances occur, the integrated control method is expected to counteract more rapidly the effect of disturbances than batch-to-batch ILC only. Lee and co-workers [25, 90, 91] proposed a batch MPC (BMPC) technique for tracking control by incorporating the capability of real-time feedback control into Q-ILC. Chin et al. [24] proposed a two-stage control framework by combining the Q-ILC and BMPC methods to separately handle the real-time disturbance and the batch-wise persisting disturbance, respectively.

However, the aforementioned integrated control strategies [24, 25, 90, 91] rely on linear time varying (LTV) models, which is known to be lacking in the extrapolative capability. Motivated by this consideration and the benefits of the integrated control strategy, a new integrated nonlinear MPC and batch-to-batch (NMPC-B2B) control strategy is proposed in this chapter. The proposed method makes use of a hybrid model consisting of the nominal first-principles model and a correction factor obtained from an updated PLS model. One major benefit of such hybrid model is the ability to harness the extrapolative capability of the first-principles model while the PLS model provides a means for simple model updating. The NMPC based on the extended predictive self-adaptive control (EPSAC) [33, 34, 70, 134] is utilized to perform online control to handle the constraints effectively while the batch-to-batch control refines the model by learning from the previous batches. Simulation studies show that the proposed control strategy results in improved constraints handling,

faster and smoother convergence compared to the standard batch-to-batch control strategy. In the next section, the batch-to-batch (B2B) control strategy adopted in this study is explained. It is followed by the discussion of the proposed NMPC-B2B control strategy. Then, simulation studies are presented to compare the performance of the B2B and NMPC-B2B control strategies in the polymorphic transformation of L-glutamic acid from the metastable α-form to the stable β-form crystals. Lastly, conclusions are given.

7.2 Batch-to-batch (B2B) control strategy

In this study, batch-to-batch control strategy based on a hybrid model consisting of a first-principles model and a PLS model is adopted. The benefit of such model lies in its ability to exploit the extrapolative power of first-principles model while the inevitable modelling error is corrected through a simple PLS model using data from previous batches. Therefore, for any process variables of interest (such as the product quality and constrained variables) at the kth sampling time of jth batch, z_k^j, its prediction can be decomposed into two factors:

$$z_k^j = z_{\mathrm{fp},k}^j + \Delta z_{\mathrm{pls},k}^j \,, \tag{7.1}$$

where $z_{\mathrm{fp},k}^j$ is obtained from the first-principles models (Chapter 3) with nominal model parameters while $\Delta z_{\mathrm{pls},k}^j$ is the correction obtained from the PLS model using

the quadratic PLS (QPLS) [165] technique.

Generally, PLS method reduces the dimension of the predictor variables $\mathbf{X} \in \mathbb{R}^{n \times n_x}$ and response variables $\mathbf{Y} \in \mathbb{R}^{n \times n_y}$, where n, n_x, n_y denote the respective numbers of datasets, input, and output variables, by projecting them to the directions that maximize the covariance between input and output variables. The decomposition of $\mathbf{X}$ and $\mathbf{Y}$ into their score and loadings matrices is shown below:

$$\mathbf{X} = \mathbf{S}\mathbf{O}^T + \mathbf{E} = \sum_{h=1}^{n_{\text{latent}}} \mathbf{s}_h \mathbf{o}_h^T + \mathbf{E}\,, \tag{7.2}$$

$$\mathbf{Y} = \mathbf{U}\mathbf{Q}^T + \mathbf{F} = \sum_{h=1}^{n_{\text{latent}}} \mathbf{u}_h \mathbf{q}_h^T + \mathbf{F}\,, \tag{7.3}$$

where $\mathbf{S} \in \mathbb{R}^{n \times n_{\text{latent}}}$ and $\mathbf{U} \in \mathbb{R}^{n \times n_{\text{latent}}}$ are the matrices of scores for $\mathbf{X}$ and $\mathbf{Y}$, respectively, $\mathbf{O} \in \mathbb{R}^{n_x \times n_{\text{latent}}}$ and $\mathbf{Q} \in \mathbb{R}^{n_y \times n_{\text{latent}}}$ are the matrices of loadings for $\mathbf{X}$ and $\mathbf{Y}$, respectively, $\mathbf{E}$ and $\mathbf{F}$ are matrices of residuals, $\mathbf{s}_h$, $\mathbf{o}_h$, $\mathbf{u}_h$ and $\mathbf{q}_h$ are the hth column of matrices $\mathbf{S}$, $\mathbf{O}$, $\mathbf{U}$ and $\mathbf{Q}$, respectively, and n_{latent} is the number of the specified latent variables used in PLS. In the conventional PLS approach, the score vectors $\mathbf{s}_h$ and $\mathbf{u}_h$ are related linearly. In contrast, they are related quadratically in the QPLS as follows:

$$\mathbf{u}_h = c_{0h} + c_{1h}\mathbf{s}_h + c_{2h}\mathbf{s}_h^2 + \varepsilon_h\,, \tag{7.4}$$

where c_{jh} is the jth regression coefficient and ε_h is the residual vector.

The first step to obtain $\Delta z^j_{\text{pls},k}$ is to prepare the matrices $\mathbf{X}$ and $\mathbf{Y}$ from the

historical database. Each row of matrix $\mathbf{X}$ consists of the input variables at sampling time 0 to $k-1$ for a particular batch, while each row of matrix $\mathbf{Y}$ contains the deviation between real process variable and the one predicted by the first-principles model at sampling time k in the same batch. There are two common approaches to determine the number of datasets (n) kept in the database. The first approach is to keep the datasets from all past batches (n increases every batch) and the other approach is to keep only the datasets from the latest n batches (i.e. moving window approach). The second step to obtain $\Delta z^j_{\mathrm{pls},k}$ is to decompose both database matrices into their corresponding scores and loadings vectors. Subsequently, the regression coefficients are obtained by the QPLS algorithm given in the Appendix A. In this study, n_{latent} is chosen as the maximum number such that the explained variances in matrices $\mathbf{X}$ and $\mathbf{Y}$ does not exceed 99%. Finally, for a new input vector $\mathbf{x}_{\mathrm{pls}}$, the output correction term $\Delta z^j_{\mathrm{pls},k}$ can be obtained as follows:

(1) Arrange the row vector $\mathbf{x}_{\mathrm{pls}}$ in the same way as the database matrix $\mathbf{X}$.

(2) For $h = 1, \ldots, n_{\mathrm{latent}}$, calculate the contributions to the output vector $\mathbf{y}_{\mathrm{pls},h}$ as follows:

(a) Obtain the input score vector $\hat{\mathbf{s}}_h$ corresponding to the new input vector:

$$\hat{\mathbf{s}}_h = \frac{\mathbf{x}_{\mathrm{pls}}\mathbf{o}_h}{\mathbf{o}_h^T\mathbf{o}_h}\,.$$

(b) Calculate the output score vector $\hat{\mathbf{u}}_h$:

$$\hat{\mathbf{u}}_h = c_{0h} + c_{1h}\hat{\mathbf{s}}_h + c_{2h}\hat{\mathbf{s}}_h^2 \,.$$

(c) Obtain the residuals vector $\mathbf{e}$:

$$\mathbf{e} = \mathbf{x}_{\text{pls}} - \hat{\mathbf{s}}_h \mathbf{o}_h^T \,,$$

and set $\mathbf{x}_{\text{pls}} = \mathbf{e}$ for the next dimension $h = h + 1$.

(3) Calculate the output correction term:

$$\Delta z^j_{\text{pls},k} = \sum_{h=1}^{n_{\text{latent}}} \hat{\mathbf{u}}_h \mathbf{q}_h^T \,.$$

In the batch-to-batch control strategy, the objective function to be minimized before the jth batch is as follows:

$$J_{\text{B2B}} = \min_{\mathcal{U}} \; W_p \left(P - P_d\right)^2 + \Delta\mathcal{U}^T \mathbf{W}_{\Delta\mathcal{U}} \, \Delta\mathcal{U} + \mathrm{d}\mathcal{U}^T \mathbf{W}_{\mathrm{d}\mathcal{U}} \, \mathrm{d}\mathcal{U} \,, \tag{7.5}$$

where

$$\begin{aligned}
\mathcal{U} &= \left[u_0^j, u_1^j, \cdots, u_{N-1}^j\right]^T \,, \\
\boldsymbol{\Delta}\mathcal{U} &= \left[u_1^j - u_0^j, u_2^j - u_1^j, \cdots, u_{N-1}^j - u_{N-2}^j\right]^T \,, \\
\mathrm{d}\mathcal{U} &= \left[u_0^j - u_0^{j-1}, u_1^j - u_1^{j-1}, \cdots, u_{N-1}^j - u_{N-1}^{j-1}\right]^T \,,
\end{aligned}$$

and P and P_d are the predicted and desired final product quality, which corresponds to either Eq. (5.1) or (5.2), respectively, u_k^j is the input value at the kth sampling time of jth batch, N is the total number of samples in one batch, W_p is the scalar weight corresponding to the final product quality, and $\mathbf{W}_{\Delta\mathcal{U}}$ and $\mathbf{W}_{\mathrm{d}\mathcal{U}}$ are the weight matrices which penalize excessive changes in the input variable which occur within-batch and inter-batch, respectively. The above minimization problem is subject to process model and inequality constraints $\mathbf{H}(\mathcal{U}) \leq 0$. In this study, differential evolution (DE) [84, 151] technique is utilized to solve the above minimization problem before every batch.

7.3 Integrated NMPC and batch-to-batch (NMPC-B2B) control strategy

The main shortcoming of batch-to-batch control strategy lies in its open-loop nature, where the correction is not made until the next batch. As a result, its capability to handle constraints for the current batch solely depends on the accuracy of the corrected model from the previous batch. Hence, when the corrected model is still not accurate, which is likely the case in the first few batches, it is possible that the constraints will be violated when the input values are implemented to the process. If online measurement of some process variables are available, it is possible and beneficial to integrate nonlinear model predictive control (NMPC) technique into the batch-to-batch control strategy to develop the proposed integrated NMPC-B2B con-

trol strategy. As such, both control strategies will complement each other such that the NMPC can perform online control to handle the constraints effectively while the batch-to-batch control refines the model by learning from the previous batches.

In the proposed integrated control strategy, the formulation of the hybrid model remains the same, except that the definition of the matrix $\mathbf{X}$ in the PLS model includes both the input and measured variables at sampling time 0 to $k-1$. The objective function to be minimized at every sampling time is as follows:

$$J_{\mathrm{NMPC-B2B}} = \min_{\mathbf{U}} \; W_p \left(P - P_d\right)^2 + \Delta\mathbf{U}^T \, \mathbf{W}_{\Delta\mathbf{U}} \, \Delta\mathbf{U} + \mathrm{d}\mathbf{U}^T \, \mathbf{W}_{\mathrm{d}\mathbf{U}} \, \mathrm{d}\mathbf{U} \,, \quad (7.6)$$

where

$$\begin{aligned} \mathbf{U} &= \left[u_k^j, u_{k+1}^j, \cdots, u_{N-1}^j\right]^T \,, \\ \boldsymbol{\Delta}\mathbf{U} &= \left[u_k^j - u_{k-1}^j, u_{k+1}^j - u_k^j, \cdots, u_{N-1}^j - u_{N-2}^j\right]^T \,, \\ \mathrm{d}\mathbf{U} &= \left[u_k^j - u_k^{j-1}, u_{k+1}^j - u_{k+1}^{j-1}, \cdots, u_{N-1}^j - u_{N-1}^{j-1}\right]^T \,, \end{aligned}$$

and $\mathbf{W}_{\Delta\mathbf{U}}$ and $\mathbf{W}_{\mathrm{d}\mathbf{U}}$ are the weight matrices which penalize excessive changes in the the input variable which occur within-batch and inter-batch, respectively. The above minimization problem is subject to process model and inequality constraints $\mathbf{H}\left(\mathbf{U}\right) \leq 0$.

The NMPC strategy considered here is based on the EPSAC technique [33, 34, 70, 134] as described in Chapter 6. Using the representation (6.12), P, $\Delta\mathbf{U}$, and

$\mathrm{d}\mathbf{U}$ in Eq. (7.6) can be decomposed into:

$$P = P_b + \mathbf{G}_{pl}\Delta\mathbf{U}_l\,, \tag{7.7}$$

$$\Delta\mathbf{U} = \Delta\mathbf{U}_b + \Delta\mathbf{U}_l\,, \tag{7.8}$$

$$\mathrm{d}\mathbf{U} = \mathbf{U}_b + \mathbf{M}\Delta\mathbf{U}_l - \mathbf{U}_{\mathrm{prev}}\,, \tag{7.9}$$

where P_b is the product quality calculated using the hybrid model with predetermined future inputs $\mathbf{U}_b = [u_{b,k}, u_{b,k+1}, \cdots, u_{b,N-1}]^T$, $\mathbf{G}_{pl}$ is the step response coefficient matrix corresponding to the product quality, $\mathbf{U}_{\mathrm{prev}} = \left[u_k^{j-1}, u_{k+1}^{j-1}, \cdots, u_{N-1}^{j-1}\right]^T$ is the input sequence from the previous batch, $\Delta\mathbf{U}_b = [\Delta u_{b,k}, \Delta u_{b,k+1}, \cdots, \Delta u_{b,N-1}]^T$ is the change in the predetermined future inputs, and M is a lower triangular matrix with all elements equal to one. Therefore, the the minimization problem (7.6) becomes:

$$J_{\mathrm{NMPC-B2B}} = \min_{\Delta\mathbf{U}_l} \quad \Delta\mathbf{U}_l^T\Gamma\Delta\mathbf{U}_l + \psi^T\Delta\mathbf{U}_l\,, \tag{7.10}$$

where

$$\Gamma = W_p\mathbf{G}_{pl}^T\mathbf{G}_{pl} + W_{\Delta\mathbf{U}} + \mathbf{M}^T\mathbf{W}_{\mathrm{d}\mathbf{U}}\mathbf{M}\,,$$

$$\psi = 2\left[W_p\left(P_b - P_d\right)^T G_{pl} + \Delta\mathbf{U}_b^T\mathbf{W}_{\Delta\mathbf{U}} + \left(\mathbf{U}_b - \mathbf{U}_{\mathrm{prev}}\right)^T\mathbf{W}_{\mathrm{d}\mathbf{U}}\mathbf{M}\right]^T\,.$$

Similarly, the inequality constraints $\mathbf{H}(\mathbf{U})$ can be decomposed into:

$$\mathbf{H}_b + \mathbf{G}_{hl}\Delta\mathbf{U}_l \leq 0\,, \tag{7.11}$$

where $\mathbf{G}_{hl}$ is the step response coefficient matrix corresponding to the constraints and $\mathbf{H}_b$ is the constraints calculated using the hybrid model with predetermined future inputs $\mathbf{U}_b$. In this study, the soft-constraint approach [140] is utilized and the minimization problem is modified as follows:

$$\min_{\Delta\mathbf{U}_l,\epsilon} J_{\text{sc,NMPC-B2B}}\,, \tag{7.12}$$

subject to

$$\mathbf{H}_b + \mathbf{G}_{hl}\Delta\mathbf{U}_l \leq \epsilon\,, \tag{7.13}$$

$$\epsilon \geq 0\,, \tag{7.14}$$

where $J_{\text{sc,NMPC-B2B}} = J_{\text{NMPC-B2B}} + \epsilon^T\,\mathbf{W}_\epsilon\,\epsilon + \epsilon^T\,\mathbf{w}_\epsilon$, ϵ is a vector of slack variables, $\mathbf{W}_\epsilon$ is a diagonal matrix of positive weight, and $\mathbf{w}_\epsilon$ is a vector of positive element.

Hence, the solution to the modified minimization problem is as follows:

$$J^*_{\text{sc,NMPC-B2B}} = \min_{\Delta\mathbf{U}_l,\epsilon} \Delta\mathbf{U}_l^T\Gamma\Delta\mathbf{U}_l + \psi^T\Delta\mathbf{U}_l + \epsilon^T\mathbf{W}_\epsilon\epsilon + \epsilon^T\mathbf{w}_\epsilon$$

$$
\begin{aligned}
&= \min_{\Delta\mathbf{U}_l,\epsilon} \begin{bmatrix} \Delta\mathbf{U}_l^T & \epsilon^T \end{bmatrix} \begin{bmatrix} \Gamma & 0 \\ 0 & \mathbf{W}_\epsilon \end{bmatrix} \begin{bmatrix} \Delta\mathbf{U}_l \\ \epsilon \end{bmatrix} + \\
&\quad \begin{bmatrix} \psi^T & \mathbf{w}_\epsilon^T \end{bmatrix} \begin{bmatrix} \Delta\mathbf{U}_l \\ \epsilon \end{bmatrix} \\
&= \min_{\Pi} \ \Pi^T \Lambda \Pi + \tau^T \Pi \,, \qquad (7.15)
\end{aligned}
$$

where

$$
\begin{aligned}
\Pi &= \begin{bmatrix} \Delta\mathbf{U}_l^T & \epsilon^T \end{bmatrix}^T , \\
\Lambda &= \begin{bmatrix} \Gamma & 0 \\ 0 & \mathbf{W}_\epsilon \end{bmatrix} , \\
\tau &= \begin{bmatrix} \psi^T & \mathbf{w}_\epsilon^T \end{bmatrix}^T ,
\end{aligned}
$$

subject to

$$
\begin{bmatrix} \mathbf{H}_b \\ 0 \end{bmatrix} + \begin{bmatrix} \mathbf{G}_{hl} & -\mathbf{I} \\ 0 & -\mathbf{I} \end{bmatrix} \Pi \leq 0 \,. \qquad (7.16)
$$

In summary, the procedure of implementing the integrated NMPC and batch-to-batch control strategy for each batch j and sampling time k is as follows:

(1) Prepare the database matrices $\mathbf{X}$ and $\mathbf{Y}$ for the PLS model as follows:

- if $j = 1$, the database matrices $\mathbf{X}$ and $\mathbf{Y}$ for the PLS model can be obtained from the historical batch data. Alternatively, input sequences

around the nominal trajectory which is the optimal input sequence for the first-principles model are implemented to the process and the resulting measurements are used to construct the database.

- if $j > 1$, update the database matrices by including the previous batch measurements dataset into the database. In this study, the moving window approach is adopted, where the dataset from the earliest batch is removed every time a new dataset is included.

(2) Obtain $\mathbf{U}_b$ by the following method:

- if $k = 0$ and $iter = 1$, $\mathbf{U}_b$ is chosen to be the input trajectory implemented in the previous batch.
- if $k > 0$ and $iter = 1$, $\mathbf{U}_b$ is set as the $\mathbf{U}_{optimal}$ obtained in the previous sampling time of the current batch.
- if $iter > 1$, the updated $\mathbf{U}_b$ from the previous iteration is used.

where $iter$ is the iteration count.

(3) Obtain P_b and $\mathbf{H}_b$ by using $\mathbf{U}_b$ as the input to the hybrid process model. In this study, it is assumed that the constrained variables are measured. Then, the bias between the predictions and the measurements of the constrained variables at the current sampling time k are added into the future predictions. If the constrained variables are not measured, it is possible to employ state estimation such as extended Kalman filter (EKF) or unscented Kalman filter

(UKF) to estimate them.

(4) Obtain the step response coefficient matrices $\mathbf{G}_{pl}$ and $\mathbf{G}_{hl}$ by introducing step change in δu. Generally, product quality measurement of batch process is only available at the end of the batch. Consequently, the PLS correction can only be calculated for the end of batch only (i.e. for sampling time N). Therefore in order to obtain $\mathbf{G}_{pl}$, the PLS correction for the sampling time N is added to the prediction of product quality at sampling time k to $N-1$.

(5) Obtain $\Pi^* = \begin{bmatrix} \Delta \mathbf{U}_l^* & \epsilon^* \end{bmatrix}^T$ from the solution to the minimization problem (7.15) and (7.16), then update the elements of $\mathbf{U}_b$ as follows:

$$u_{b,k+j} = u_{b,k+j} + \sum_{i=0}^{j} \Delta u_{l,k+i}\,,$$

where $j = 0, \ldots, N-1+k$.

(6) If $err = \left\| \begin{bmatrix} \mathbf{G}_{pl} \\ \mathbf{G}_{hl} \end{bmatrix} \Delta \mathbf{U}_l^* \right\|$ is greater than a specified tolerance (1×10^{-4} is used in this study), $iter = iter + 1$ and repeat from step (2). Otherwise, set $\mathbf{U}_{optimal} = \mathbf{U}_b$ and implement the first element of $\mathbf{U}_{optimal}$ to the process.

(7) If the end of the current batch is reached, repeat from step (1) and go to the next batch.

7.4 Simulation results and discussion

In the polymorphic transformation process, both α- and β-form crystals are seeded according to Gaussian distribution with parameter values given in Table 5.1. The initial solute concentration C_0 and maximum final solute concentration $C_{\max}(t_f)$ are 20 $\mathrm{g/kg}$. The batch time t_f is 3 hours and the sampling time is ten minutes.

7.4.1 Description of specific control implementations

The optimization of two different product qualities, $P_1 = (\mu_{\beta,3})_{t=t_f}$ in Eq. (5.1) and $P_2 = \left(\frac{\mu_{\beta,3}^{\text{nucl}}}{\mu_{\beta,3}^{\text{seed}}}\right)_{t=t_f}$ in Eq. (5.2), are considered, which from here onwards will be called objective J_1 and objective J_2, respectively. It is assumed that the process is subject to two cases of parameter perturbations given in Table 5.2. Note that only Cases 2 and 3 are considered, since batch-to-batch adjustment is not required when there is no plant-model mismatch (Case 1). The tuning parameters for the B2B and NMPC-B2B control strategies for both objectives are given in Tables 7.1 and 7.2, respectively. For all cases and objectives, the initial database (i.e. for the first batch) utilized for the PLS model comprises historical operating data from ten batches. These include temperature trajectories around the nominal trajectory (see the solid lines in Figures 7.1 to 7.4) obtained by optimizing the nominal first-principles model, the corresponding deviation between the measured concentration and the predicted concentration by the first-principles model, and the deviation values in final product quality. For the subsequent batches, moving window approach

is adopted to update the database, where the latest 15 batches are kept in the database. For all cases and objectives, the optimal results as indicated by the dashed line in Figures 7.5 to 7.10 are obtained by applying the temperature control strategy, where the temperature-time trajectory is parameterized as a first-order spline with eighteen time intervals, to the first-principles model with the set of parameters treated as known.

7.4.2 Comparison results and discussion

For the first control objective J_1, the respective concentration and temperature trajectories obtained for the B2B and NMPC-B2B control strategies applied to Case 2 are given by Figures 7.5 and 7.6. As can be seen from these figures, both control strategies produce solutions which converge to the optimal one gradually and result in temperature and concentration trajectories very close to the corresponding optimal ones at batch 20. Figure 7.7 compares the P_1 values obtained by both control strategies for batches 1 to 20. It is clear that not only the P_1 values obtained by the NMPC-B2B control strategy converge at a faster rate than those obtained by the B2B control strategy, but also the former gives a smoother convergence while a slight oscillation can be observed in batches 13 to 18 for the B2B control strategy. Furthermore, the NMPC-B2B control strategy is able to satisfy all the constraints for every batch, while the B2B control strategy violates one of the constraints during batches 5 to 8 as shown in Figure 7.7. For Case 3, the resulting concentration and temperature trajectories for both control strategies in batches 1, 7, 14, and 20

are shown in Figures 7.8 and 7.9, where the convergence of both control strategies to the corresponding optimal one is illustrated. The trend of P_1 values as given in Figure 7.10 shows similar observation to the Case 2, meaning that the NMPC-B2B control strategy results in a faster convergence and satisfies all constraints, while the B2B control strategy violates one of the constraints in batch 2.

For control objective J_2, Figures 7.11 and 7.12 show the concentration and temperature trajectories for B2B and NMPC-B2B control strategies applied to Case 2, respectively. Although there is a slight difference between the temperature trajectories obtained by both control strategies and the optimal temperature trajectory in batch 20, the corresponding P_2 values (Figure 7.13) obtained by both control strategies are reasonably close to the optimal one (within 0.2%). This phenomenon also indicates that the P_2 value is less sensitive to the changes in temperature around the optimal temperature trajectory. Again, it is observed that the NMPC-B2B control strategy converges at a faster rate than the B2B control strategy. Likewise, Figure 7.14 shows that the P_2 values in batch 20 obtained by the two control strategies for Case 3 are very close to the optimal one (within 0.1%), despite the difference in the corresponding concentration and temperature trajectories produced by both control strategies and the optimal ones (Figures 7.15 and 7.16). In addition, the convergence of the P_2 values obtained by the NMPC-B2B control strategy is much faster and smoother than that obtained by the B2B control strategy.

The P_1 and P_2 values obtained by the B2B and NMPC-B2B control strategies at batch 20 and those obtained by NMPC strategy developed in Chapter 6 are tab-

ulated in Tables 7.3 and 7.4, respectively. It can be seen that the P_1 and P_2 values obtained by the B2B and NMPC-B2B control strategies are very close to each other and comparable to the true optimal values for all cases. Furthermore, through the learning process, these control strategies outperform the NMPC strategy, except for Case 3 of objective J_1, where the NMPC strategy obtains a marginally better P_1 value (0.1% better than that obtained by the NMPC-B2B strategy).

Table 7.1: Tuning parameters for the B2B control strategy.

Values for objective J_1	Values for objective J_2
$W_{p,1} = 1$	$W_{p,2} = 1$
$(\mathbf{W}_{\Delta\mathcal{U},1})_{i,i}{}^{\dagger} = 2\left[1 + 15\left(i-1\right)\right] \times 10^{-5}$	$(\mathbf{W}_{\Delta\mathcal{U},2})_{i,i}{}^{\dagger} = 3\left[1 + 0.5\left(i-1\right)\right] \times 10^{-5}$
$\mathbf{W}_{\mathrm{d}\mathcal{U},1} = 3 \times 10^{-5} I$	$\mathbf{W}_{\mathrm{d}\mathcal{U},2} = 5 \times 10^{-6} I$
$\mathbf{W}_{\epsilon} = 10 I$	$\mathbf{W}_{\epsilon} = 10 I$
$\mathbf{w}_{\epsilon} = 10\left[1, 1, \ldots, 1\right]^T$	$\mathbf{w}_{\epsilon} = 10\left[1, 1, \ldots, 1\right]^T$

†The diagonal elements of matrices $\mathbf{W}_{\Delta\mathcal{U},1}$ and $\mathbf{W}_{\Delta\mathcal{U},2}$, where $i = 1, \ldots, N$.

Table 7.2: Tuning parameters for the NMPC-B2B control strategy.

Values for objective J_1	Values for objective J_2
$W_{p,1} = 1$	$W_{p,2} = 1$
$(\mathbf{W}_{\Delta\mathbf{U},1})_{i,i}{}^{\dagger} = \left[1 + 15\left(i-1\right)\right] \times 10^{-5}$	$(\mathbf{W}_{\Delta\mathbf{U},2})_{i,i}{}^{\dagger} = 9\left[1 + 0.7\left(i-1\right)\right] \times 10^{-5}$
$\mathbf{W}_{\mathrm{d}\mathbf{U},1} = 1.5 \times 10^{-5} I$	$\mathbf{W}_{\mathrm{d}\mathbf{U},2} = 6 \times 10^{-6} I$
$\mathbf{W}_{\epsilon} = 10 I$	$\mathbf{W}_{\epsilon} = 10 I$
$\mathbf{w}_{\epsilon} = 10\left[1, 1, \ldots, 1\right]^T$	$\mathbf{w}_{\epsilon} = 10\left[1, 1, \ldots, 1\right]^T$

†The diagonal elements of matrices $\mathbf{W}_{\Delta\mathbf{U},1}$ and $\mathbf{W}_{\Delta\mathbf{U},2}$, where $i = 1, \ldots, N-k$.

Table 7.3: Values of the control objective P_1 obtained for the Cases 2 and 3 in Table 5.2.

Cases	NMPC	B2B	NMPC-B2B	optimal
2	0.4031	0.4194 (after batch 11)	0.4194 (after batch 8)	0.4195
3	0.2666	0.2662 (after batch 6)	0.2663 (after batch 3)	0.2667

Table 7.4: Values of the control objective P_2 obtained for the Cases 2 and 3 in Table 5.2.

Cases	NMPC	B2B	NMPC-B2B	optimal
2	0.0049	0.0046 (after batch 9)	0.0046 (after batch 7)	0.0046
3	0.0679	0.0659 (after batch 20)	0.0659 (after batch 10)	0.0659

7.5 Conclusions

An integrated nonlinear predictive control and batch-to-batch (NMPC-B2B) control strategy utilizing a hybrid model was developed for batch polymorphic crystallization processes. The performance of the proposed control strategy to optimize two control objectives P_1 and P_2 is evaluated under two cases of plant-model mismatch. Simulation results show that the NMPC-B2B control strategy produces better performance compared to the standard B2B control strategy for all cases and objectives considered. Beside being able to satisfy all the constraints, the convergence of the product qualities obtained by the NMPC-B2B control strategy is always faster and smoother than that obtained by the B2B control strategy. Compared to the NMPC strategy developed in the previous chapter, both B2B and NMPC-B2B control strategies obtain better product quality values (except for Case 3 of objective J_1). Hence, it is verified that through the learning process, both B2B and NMPC-B2B control strategies are more advantageous to be employed to address the plant-model mismatch in an effective manner.

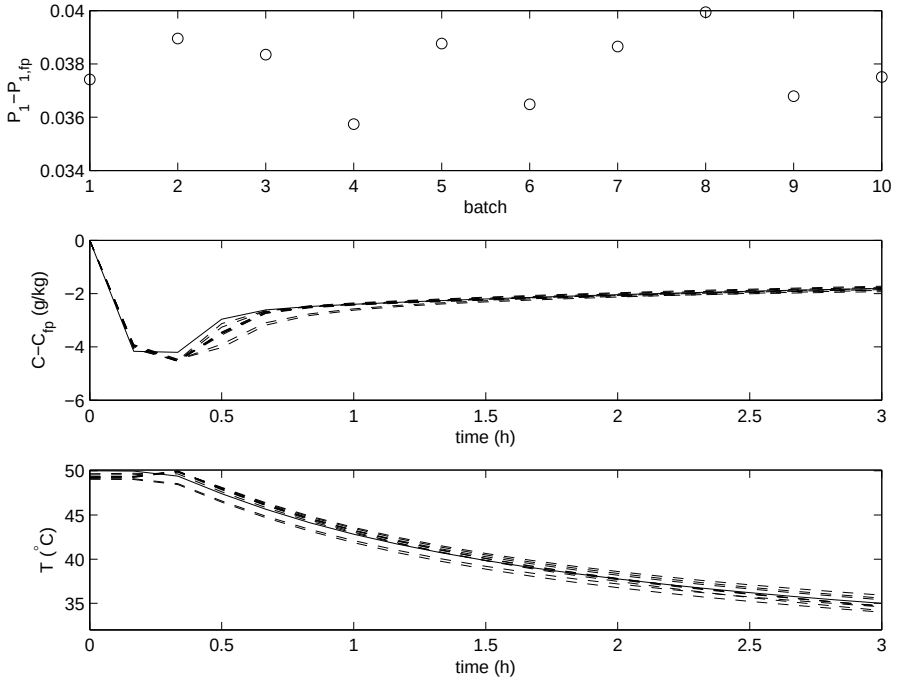


Figure 7.1: Database employed for Case 2 and objective J_1 in B2B and NMPC-B2B control strategies.

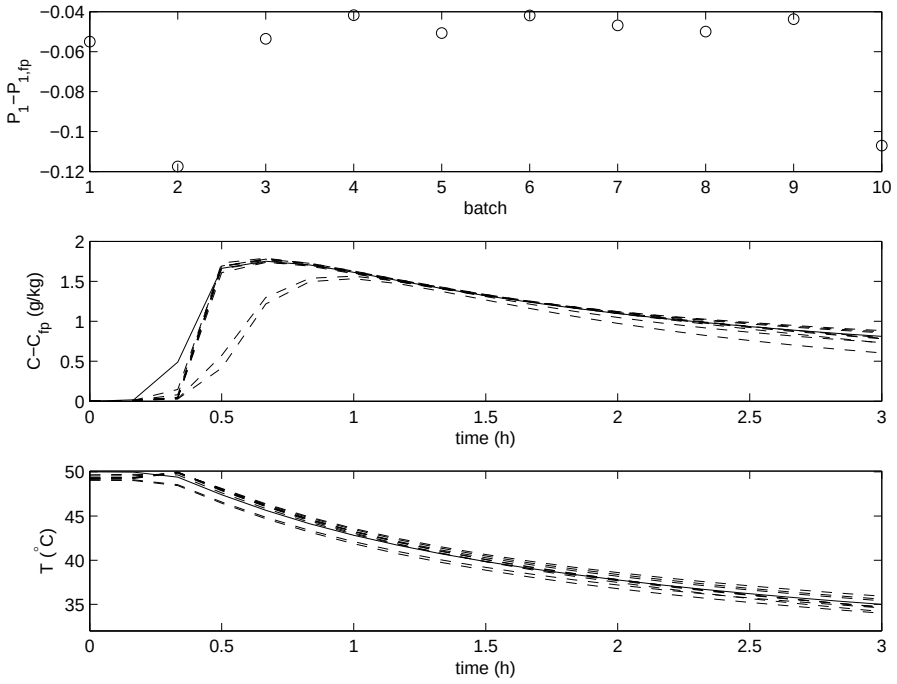


Figure 7.2: Database employed for Case 3 and objective J_1 in B2B and NMPC-B2B control strategies.

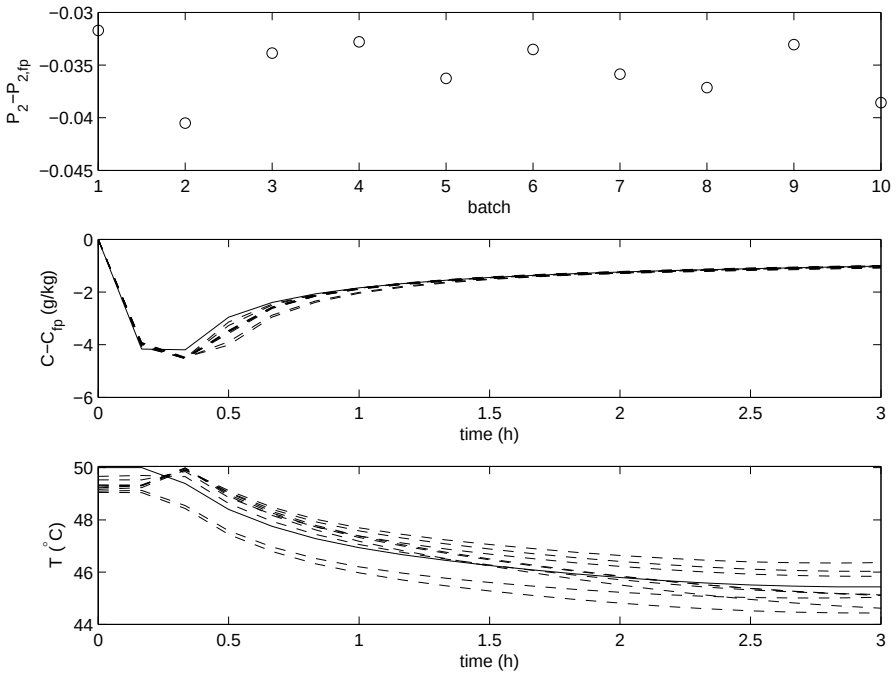


Figure 7.3: Database employed for Case 2 and objective J_2 in B2B and NMPC-B2B control strategies.

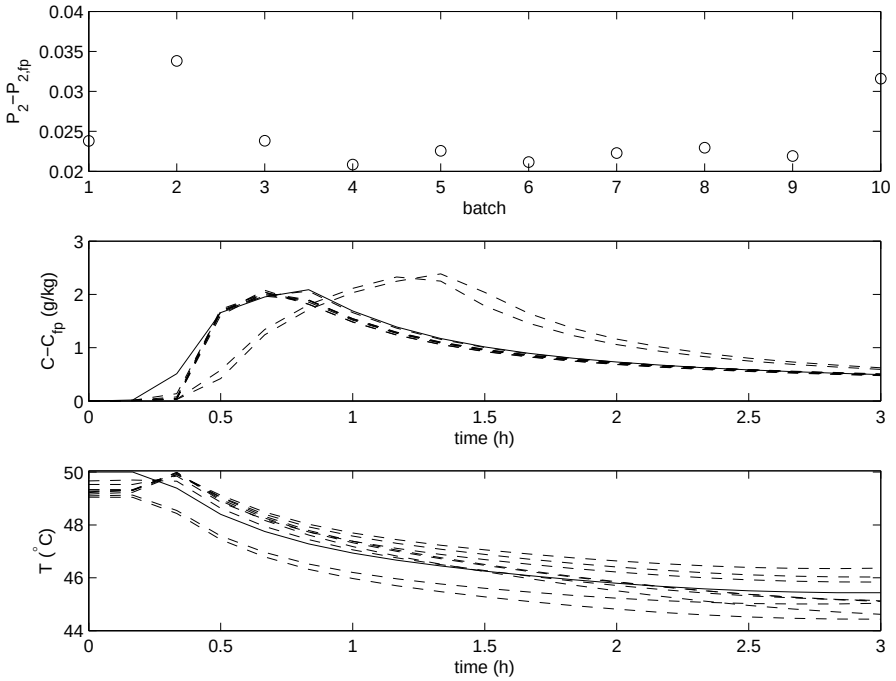


Figure 7.4: Database employed for Case 3 and objective J_2 in B2B and NMPC-B2B control strategies.

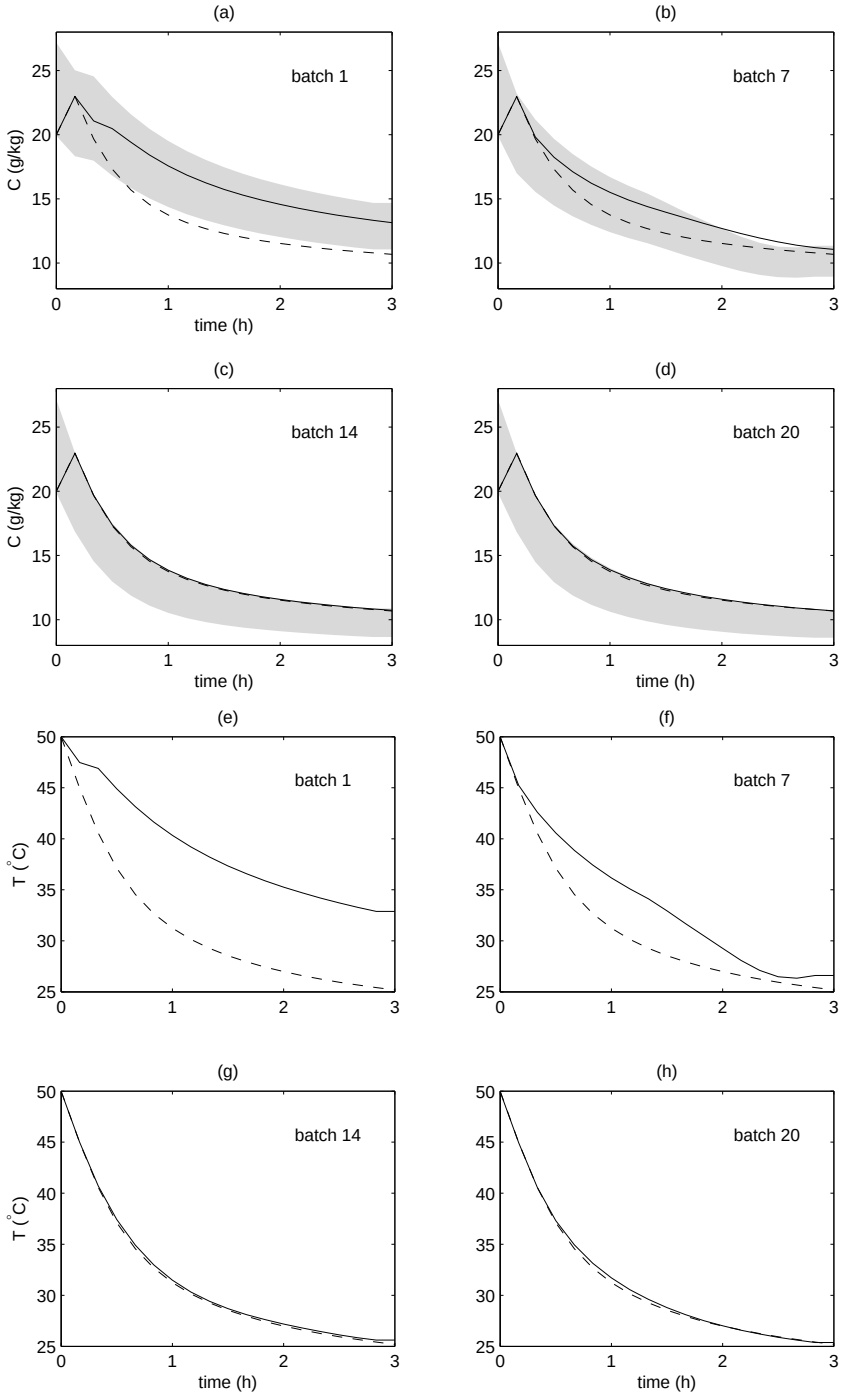


Figure 7.5: Result of B2B control strategy for Case 2 and objective J_1: (a) to (d) are the concentration trajectories and the shaded region shows the constraints on the concentration; (e) to (h) are the temperature trajectories. Solid line: B2B control, dashed line: optimal control.

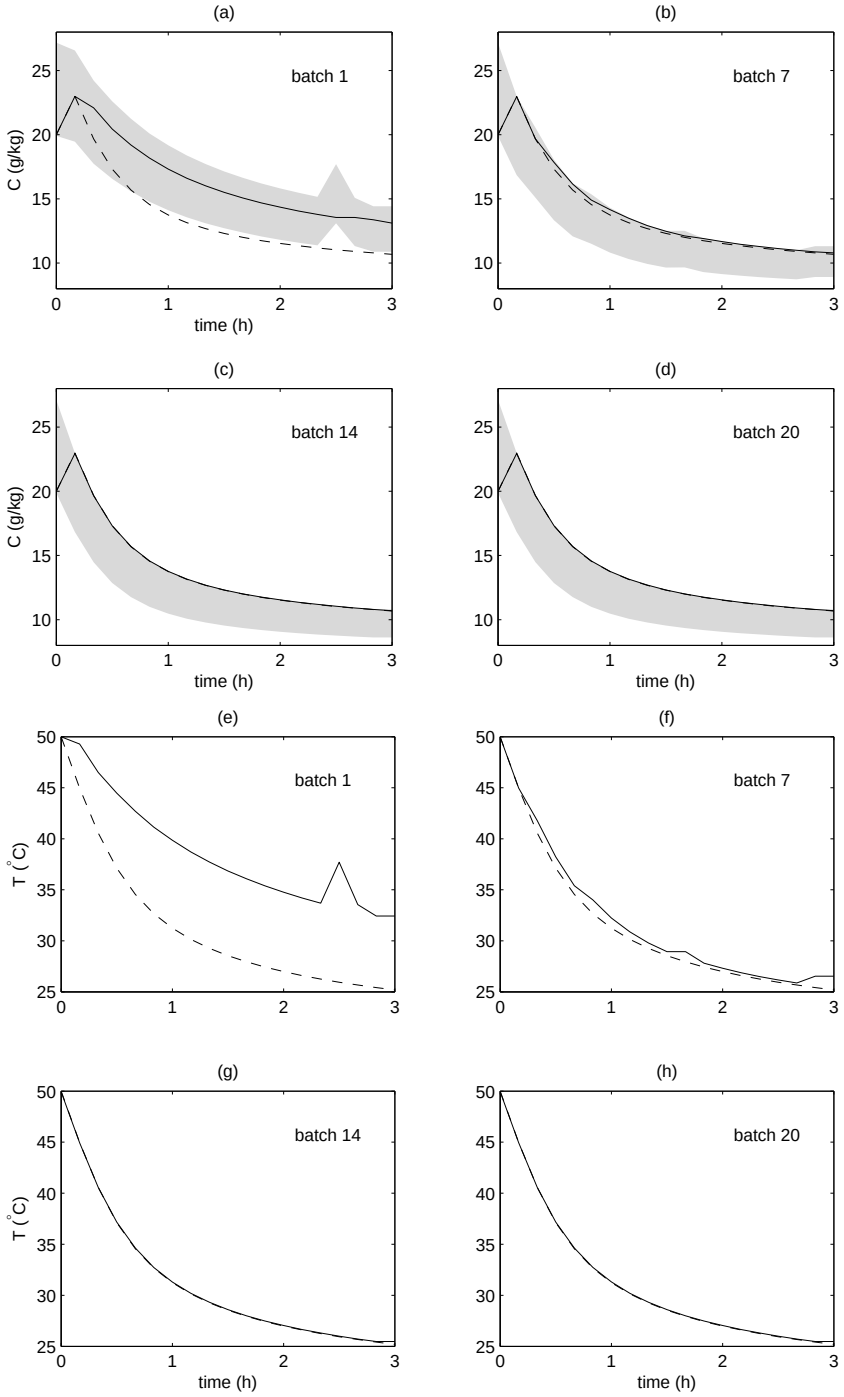


Figure 7.6: Result of NMPC-B2B control strategy for Case 2 and objective J_1: (a) to (d) are the concentration trajectories and the shaded region shows the constraints on the concentration; (e) to (h) are the temperature trajectories. Solid line: NMPC-B2B control, dashed line: optimal control.

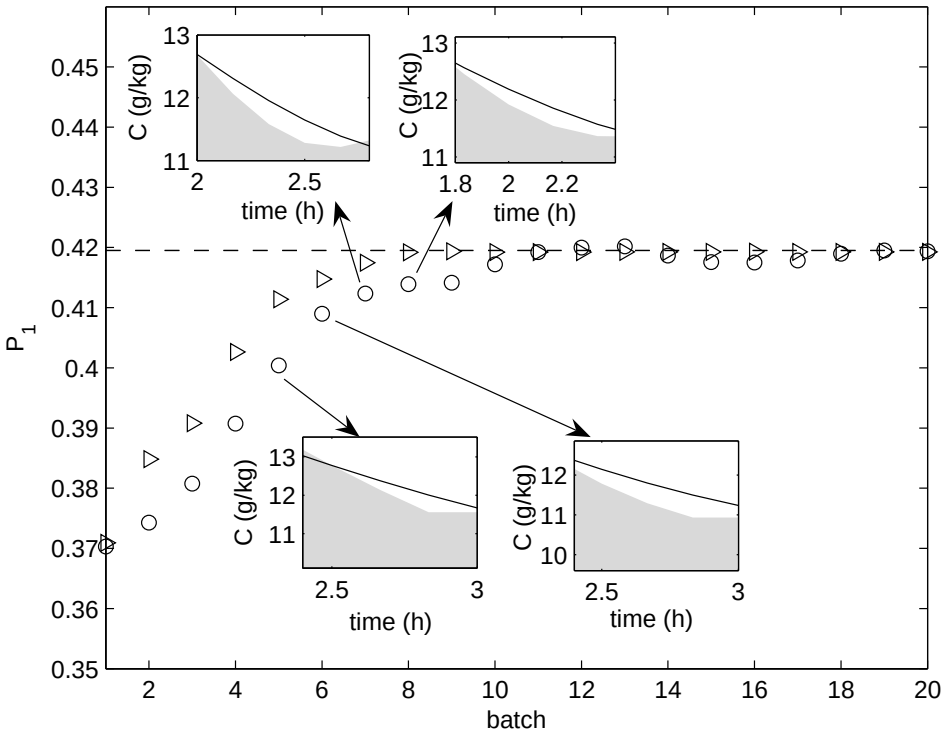


Figure 7.7: Comparison of P_1 values obtained by the B2B ($\circ$) and NMPC-B2B (Δ) control strategies for Case 2. The insets show the constraints violation for B2B control strategy in batches 5 to 8.

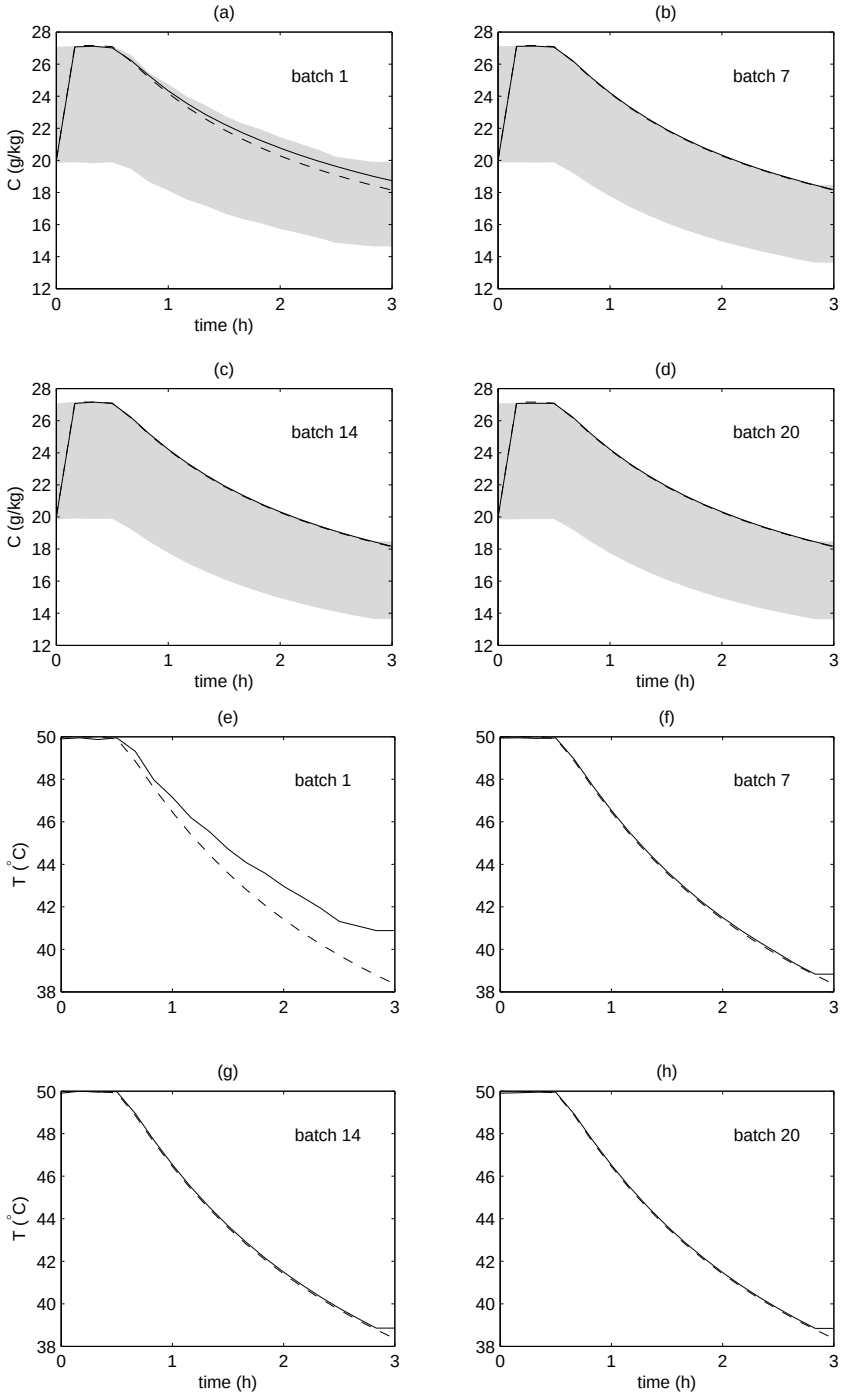


Figure 7.8: Result of B2B control strategy for Case 3 and objective J_1: (a) to (d) are the concentration trajectories and the shaded region shows the constraints on the concentration; (e) to (h) are the temperature trajectories. Solid line: B2B control, dashed line: optimal control.

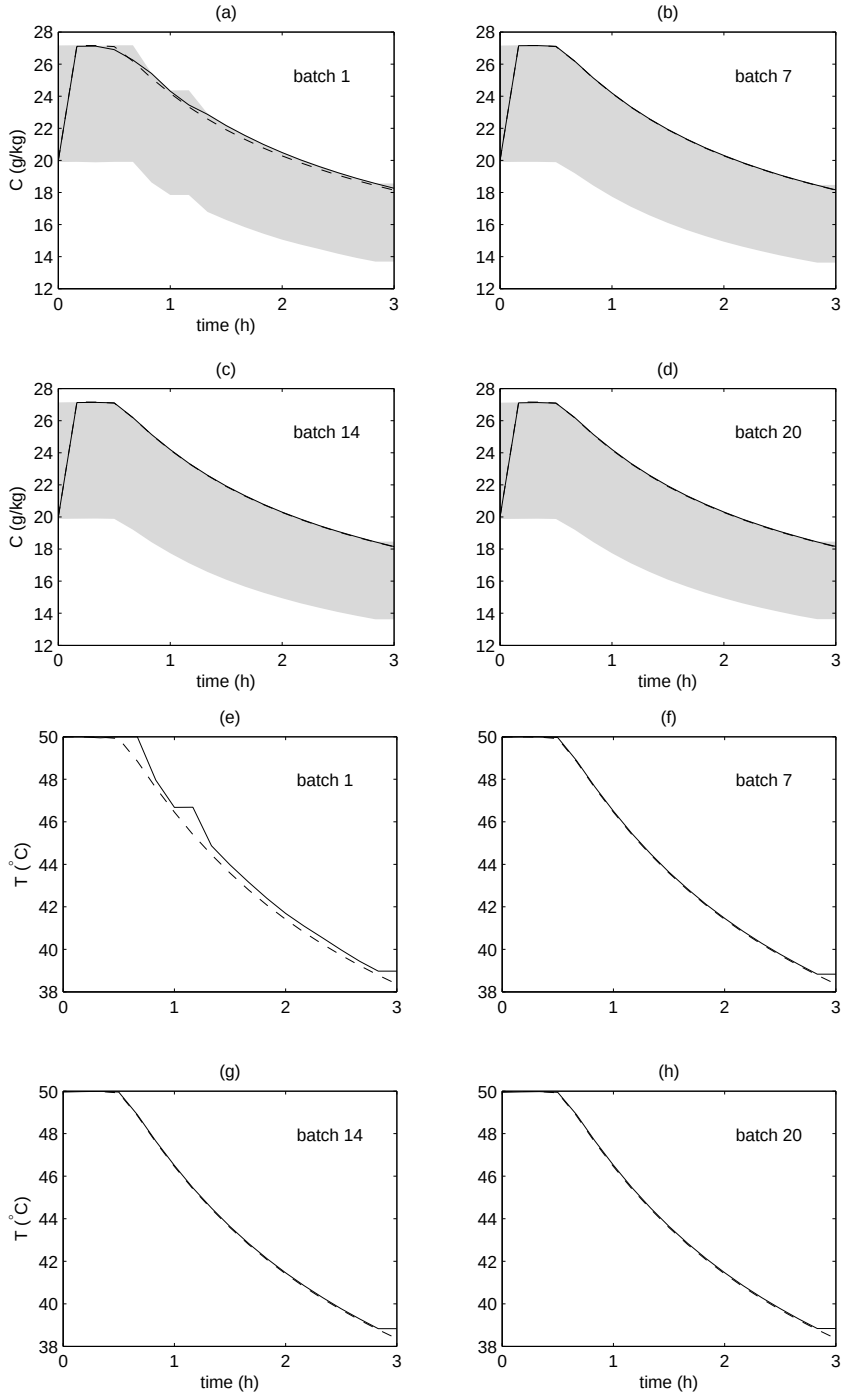


Figure 7.9: Result of NMPC-B2B control strategy for Case 3 and objective J_1: (a) to (d) are the concentration trajectories and the shaded region shows the constraints on the concentration; (e) to (h) are the temperature trajectories. Solid line: NMPC-B2B control, dashed line: optimal control.

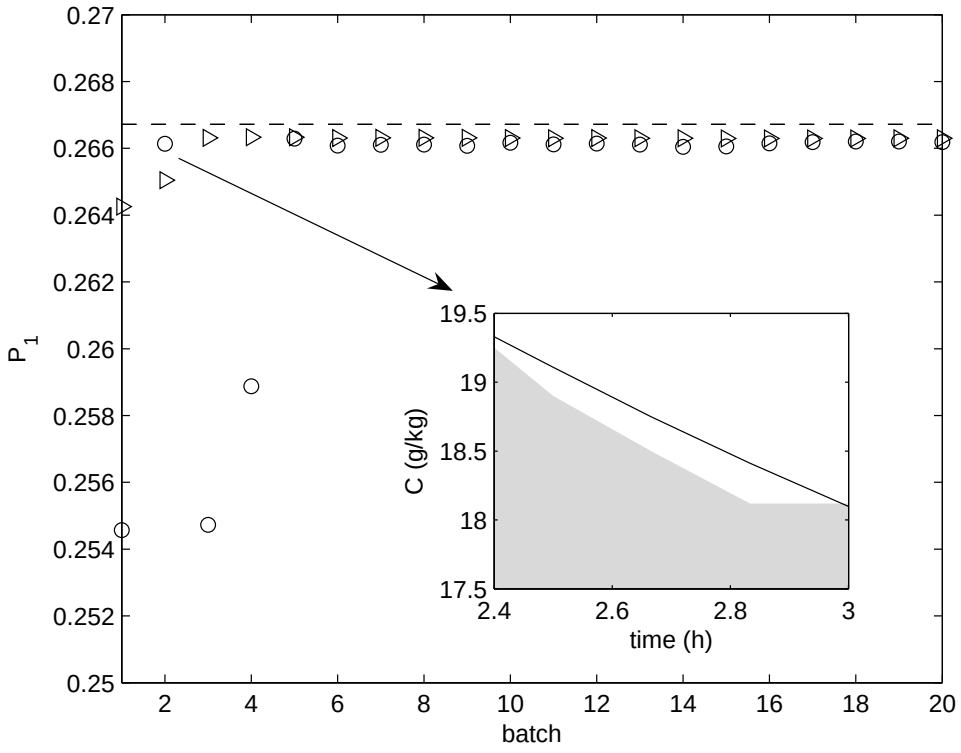


Figure 7.10: Comparison of P_1 values obtained by the B2B ($\circ$) and NMPC-B2B (Δ) control strategies for Case 3. The inset shows the constraints violation for B2B control strategy in batch 2.

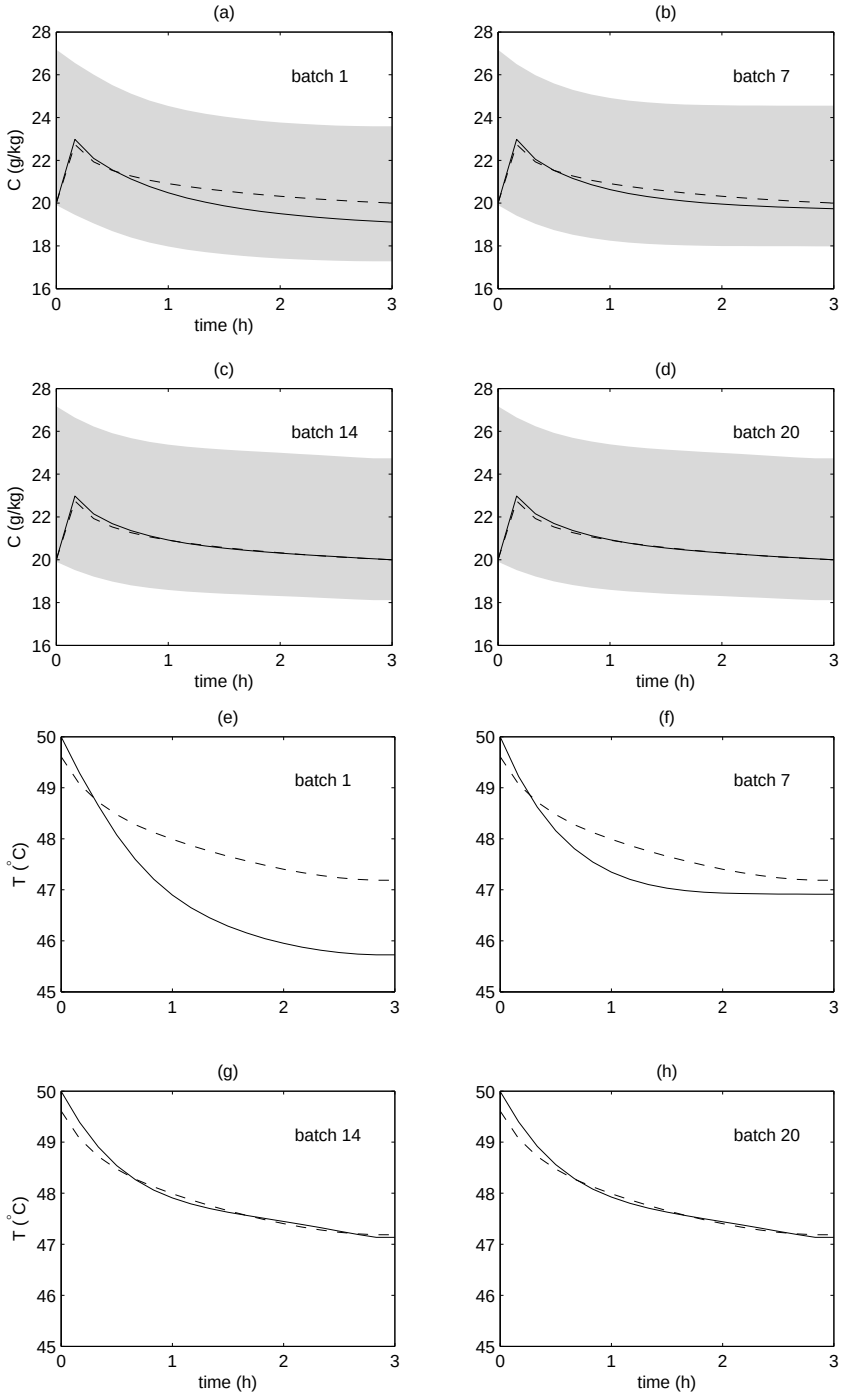


Figure 7.11: Result of B2B control strategy for Case 2 and objective J_2: (a) to (d) are the concentration trajectories and the shaded region shows the constraints on the concentration; (e) to (h) are the temperature trajectories. Solid line: B2B control, dashed line: optimal control.

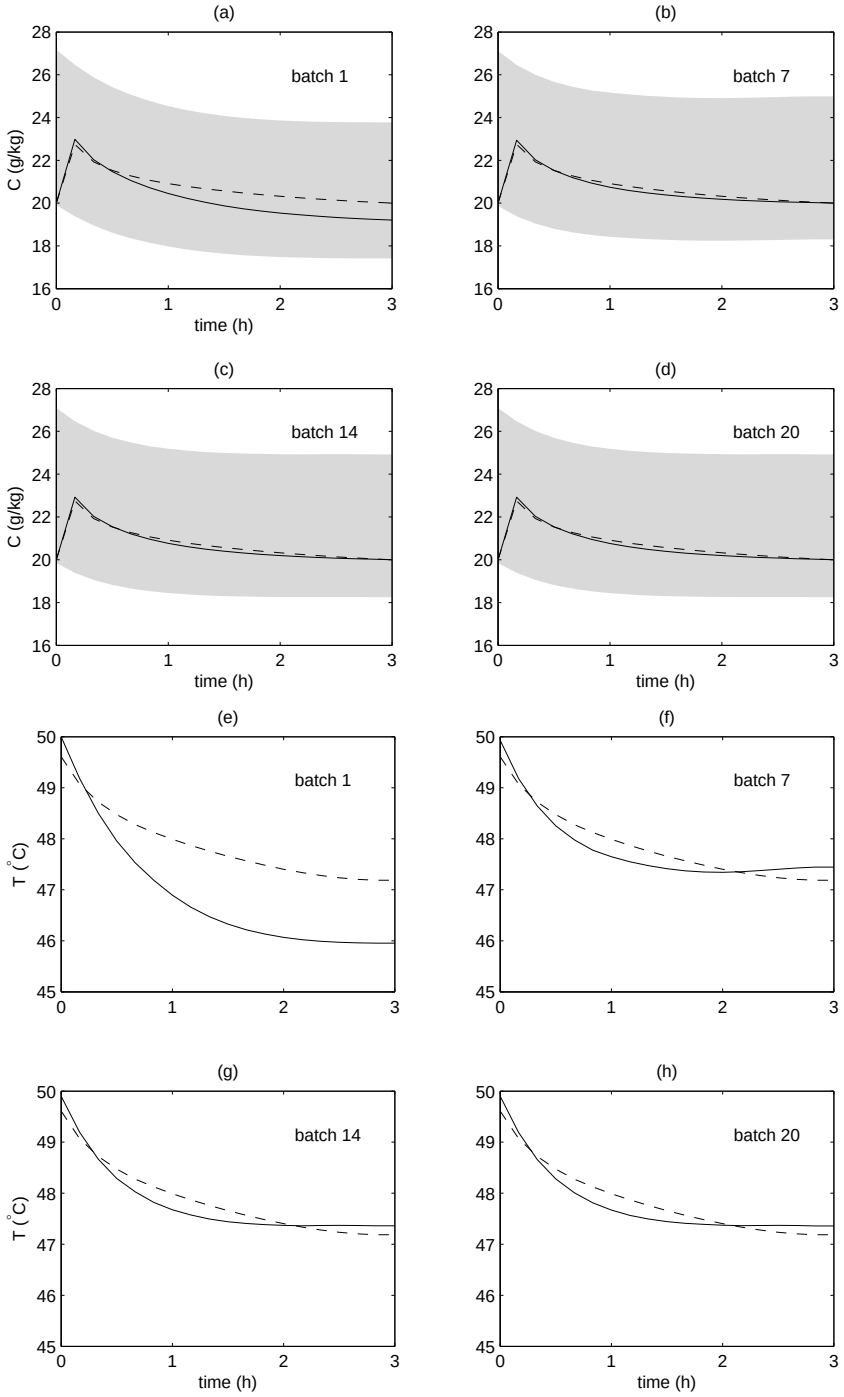


Figure 7.12: Result of NMPC-B2B control strategy for Case 2 and objective J_2: (a) to (d) are the concentration trajectories and the shaded region shows the constraints on the concentration; (e) to (h) are the temperature trajectories. Solid line: NMPC-B2B control, dashed line: optimal control.

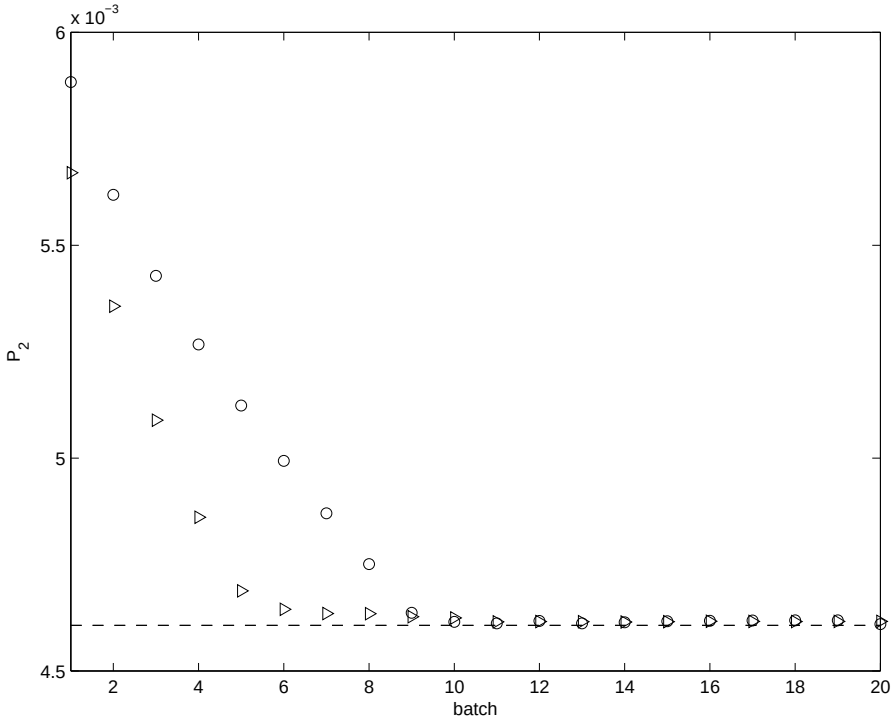


Figure 7.13: Comparison of P_2 values obtained by the B2B ($\circ$) and NMPC-B2B (Δ) control strategies for Case 2.

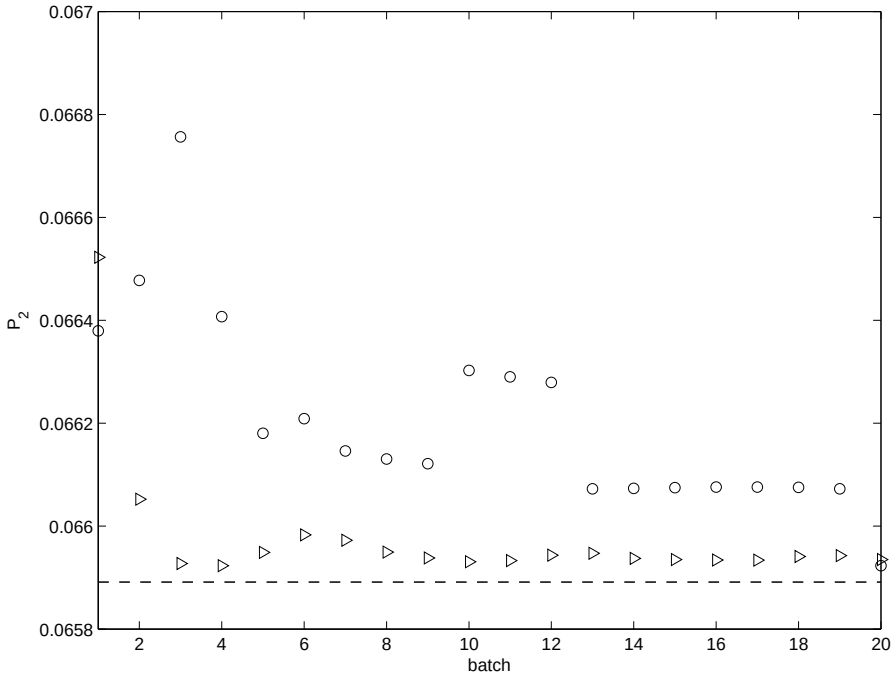


Figure 7.14: Comparison of P_2 values obtained by the B2B ($\circ$) and NMPC-B2B (Δ) control strategies for Case 3.

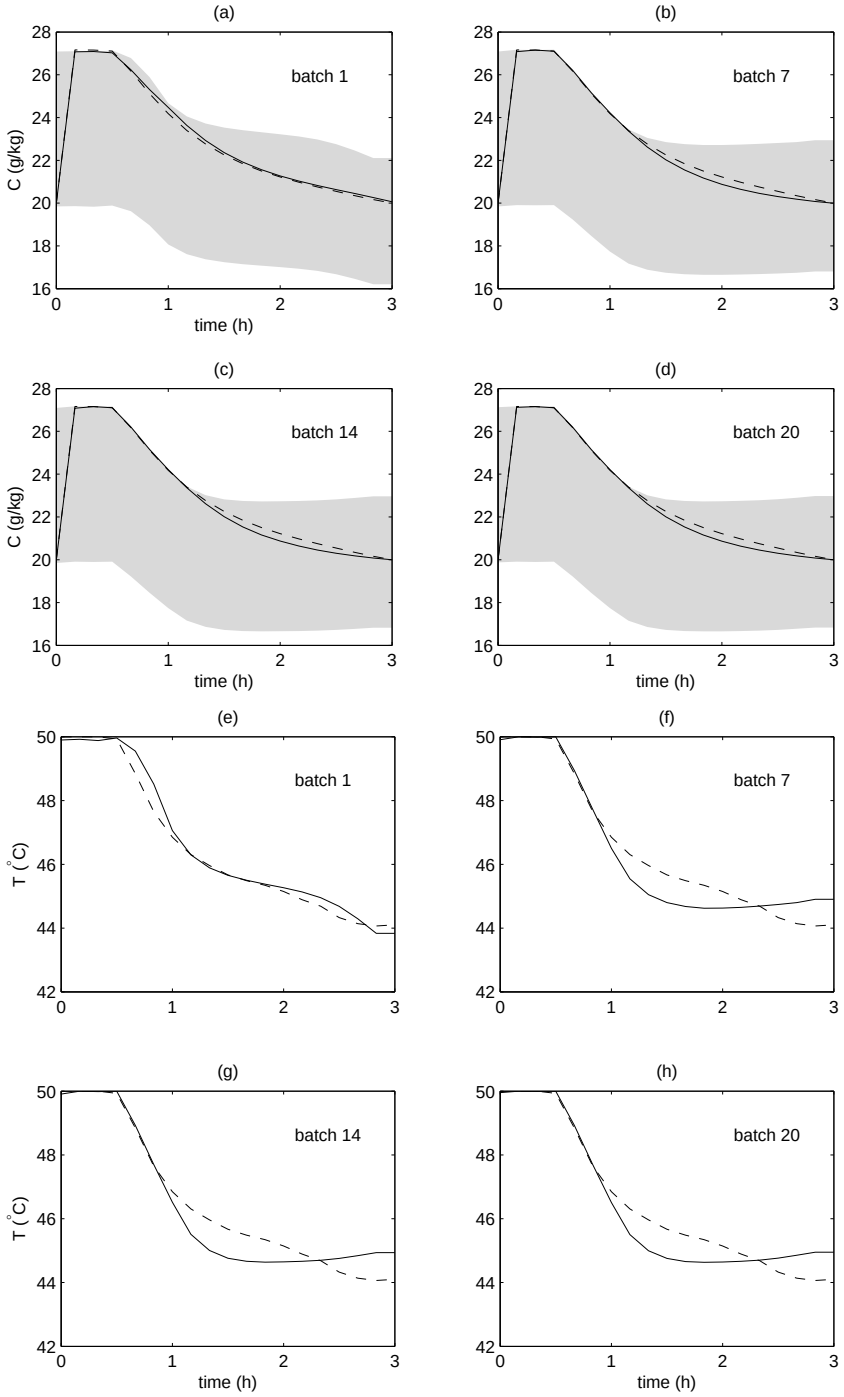


Figure 7.15: Result of B2B control strategy for Case 3 and objective J_2: (a) to (d) are the concentration trajectories and the shaded region shows the constraints on the concentration; (e) to (h) are the temperature trajectories. Solid line: B2B control, dashed line: optimal control.

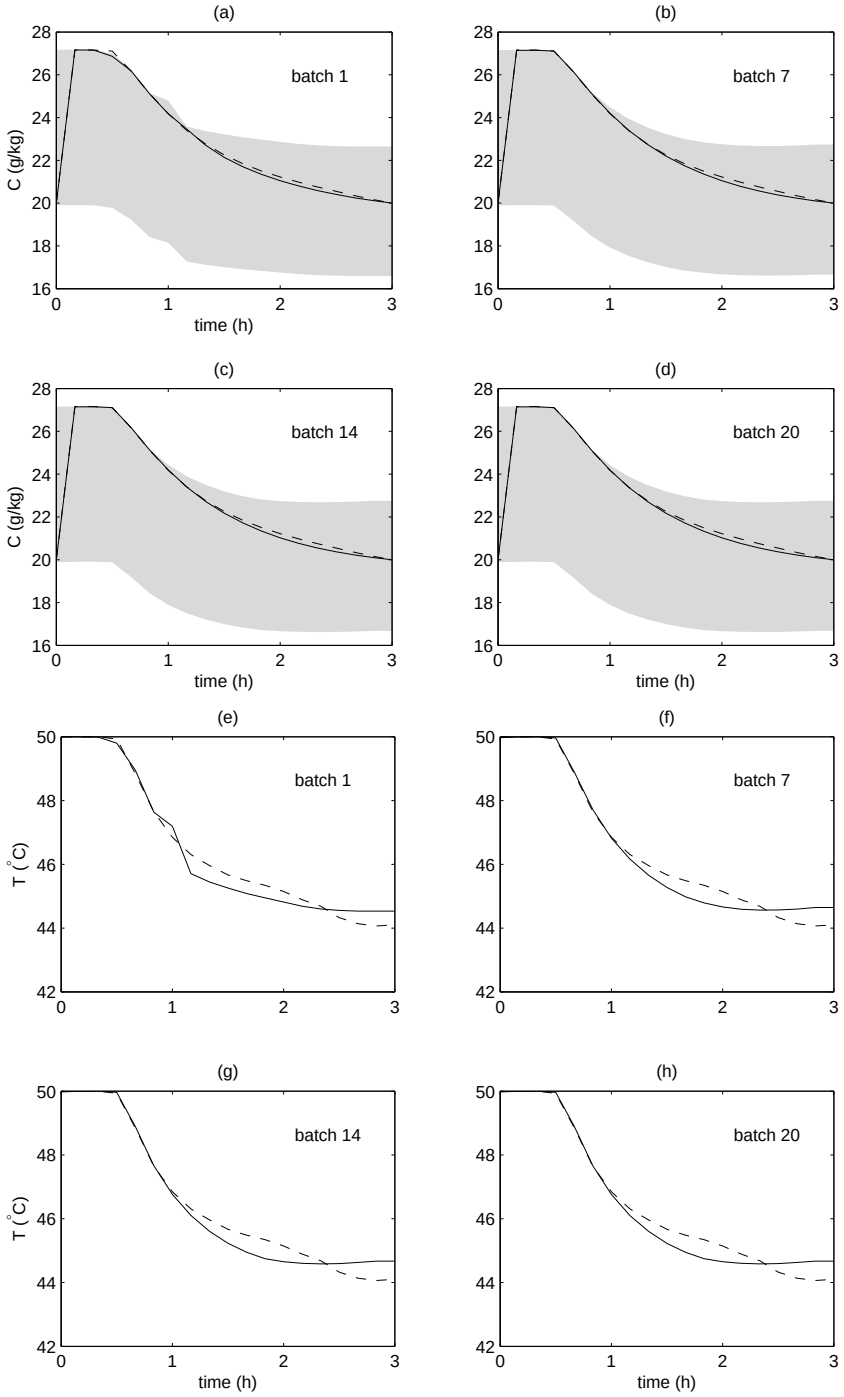


Figure 7.16: Result of NMPC-B2B control strategy for Case 3 and objective J_2: (a) to (d) are the concentration trajectories and the shaded region shows the constraints on the concentration; (e) to (h) are the temperature trajectories. Solid line: NMPC-B2B control, dashed line: optimal control.

Chapter 8

Conclusions and Future Work

8.1 Conclusions

Encouraged by the importance of polymorphism in pharmaceutical industries, this thesis investigated the modelling, simulation, and control of polymorphic crystallization of L-glutamic acid, which consists of the metastable α-form and the stable β-form crystals.

In Chapter 3, a kinetic model of L-glutamic acid polymorphic crystallization is developed from batch experiments with in-situ measurements including attenuated total reflection Fourier transform infrared (ATR-FTIR) spectroscopy to infer the solute concentration and focused beam reflectance measurement (FBRM) which provides crystal size information. The developed kinetic model appears to be the first to include all of the transformation kinetic parameters including dependence on the temperature, compared to past studies on the modelling of L-Glutamic acid

crystallization [115, 139]. The model parameters are determined using Bayesian inference instead of the standard weighted least squares methods, as such prior knowledge can be included in the statistical analysis. In addition to providing point estimates of the kinetic parameters, the Bayesian inference approach is able to determine a detailed marginal probability distribution for each parameter. The marginal probability distributions of the parameters can give practitioners insight regarding the parameter uncertainties and are of significant value to develop robust control strategies for the crystallization process [114].

In Chapter 4, numerical simulations of the polymorphic crystallization of L-glutamic acid using high-order WENO methods are developed. The performance of three WENO methods: Liu et al's version of WENO (LOCWENO) [99], Jiang and Shu's version of WENO with Henrick mapping (JSHWENO) [61, 72], and the weighted power ENO method (Wpower-ENO) [141] are compared to the high resolution (HR) finite volume method and a second-order finite difference (FD2) method. From simulation results, it is shown that the three WENO methods outperform the HR and FD2 methods in terms of computational efficiency, with LOCWENO and JSHWENO methods having the highest overall efficiency.

In Chapters 5 to 7, control strategies for the polymorphic transformation of L-glutamic acid from the metastable α-form to the stable β-form crystals, where two types of objective are considered. The first objective is to maximize the mass of β-form crystals, which is equivalent to maximizing the third-order moment or the yield of β-form crystals, whereas the second objective is to minimize the ratio of

the nucleated crystal mass to the seed crystal mass of β-form crystals.

Specifically, Chapter 5 discusses two popular control strategies in non-polymorphic crystallization, the T-control and C-control strategies. Simulation results show that T-control is in general sensitive to parameter perturbations, which is in accordance with the observation for non-polymorphic crystallization. On the other hand, the C-control strategy performs very robustly for both objectives, but long batch times may be required.

Since the method of moments which is heavily used in the previous MPC control algorithms developed for non-polymorphic crystallization processes [35, 79, 113, 131, 155] does not apply for the polymorphic transformation processes, the full PDEs of the population balance model need to be solved. Consequently, the computation time required increases considerably which prohibits the straightforward application of nonlinear programming. In Chapter 6, an efficient nonlinear predictive control (NMPC) strategy based on the extended predictive self-adaptive control (EPSAC) is developed. The resulting NMPC strategy only requires a standard quadratic programming, which increases computational efficiency considerably compared to the nonlinear programming counterpart. Compared to the T-control, C-control, and quadratic dynamic matrix control with successive linearization (SL-QDMC), the NMPC strategy shows good overall robustness for two different control objectives, which were both within 7% of their optimal values, while satisfying all constraints on manipulated and state variables within the specified batch time.

Finally, realizing that batch processes are repetitive in nature, an integrated nonlinear predictive control and batch-to-batch (NMPC-B2B) control strategy utilizing a hybrid model is developed in Chapter 7. The hybrid model comprising the nominal first-principles model and a correction factor based on an updated partial least squares (PLS) model is utilized to predict the process variables and final product quality. In the proposed NMPC-B2B control strategy, the NMPC performs online control to handle the constraints effectively while the batch-to-batch control refines the model by learning from the previous batches. Simulation studies show that the proposed NMPC-B2B control strategy produces faster and smoother convergence and satisfies all the state constraints, compared to the standard B2B control strategy. Further observations suggest that the learning process in both B2B and NMPC-B2B control strategies counteracts the plant-model mismatch effectively after several batches.

8.2 Suggestions for future work

There are few suggestions that warrant further investigation, which are summarized below.

Firstly, distributional uncertainty analysis can be carried out based on the polymorphic crystallization model developed in this thesis. Comprehensive uncertainty analysis of mechanistic models is important to quantify the influence of parameter uncertainties on the process states and outputs. This quantification eventually can be

used to design efficient schemes for model or data refinement by deciding whether more laboratory experiments are needed to provide better parameter estimates or whether other model structure should be selected. In addition, the distributional uncertainty analysis together with the posterior distribution of the model parameters provided in this thesis can be used to develop a robust T-control strategy for the polymorphic crystallization process.

Secondly, the models developed in this thesis and other literature [115, 139] assume perfect mixing. It is possible that this assumption will not be satisfied for industrial size crystallizer. Therefore, to consider the effect of hydrodynamics of the vessel which accounts for non-ideal flow behavior, compartment modelling technique can be utilized. Basically, the compartment modelling technique divides the crystallizer into few sections with different degree of mixing, in order to simulate the mixing imperfection.

Thirdly, data-based modelling techniques like neural networks can be investigated to model the polymorphic crystallization instead of the first-principles model. The benefits of modelling using data-based technique include the much less engineering effort required as compared to the first-principles modelling technique and the possibility to perform online model updating. Once data-based model can be obtained, various control strategies can be investigated based on this modelling technique and their performance can be compared to the ones developed in this thesis.

Fourthly, in the last decades the salting-out technique has drawn more attention

and has been more frequently seen as a valid alternative to cooling and evaporation. The antisolvent addition reduces the solubility of the product compound in the original solvent thereby facilitating supersaturation generation. This method is attractive since it can lead to significant savings in energy consumption and operation costs in comparison to the conventional techniques. Furthermore it can be seen as an alternative methodology whereby the limited temperature stability of the solid product precludes the use of evaporation, as in the case of pharmaceuticals and biochemicals, or when, because of the weak temperature dependence of the solute solubility, it is not possible to use the cooling techniques, such as for sodium chloride in water. Therefore, it is worthwhile to investigate the modelling, simulation, and control of antisolvent crystallization.

Appendix A

Quadratic Partial Least Squares

The following describes the quadratic partial least squares (QPLS) procedure [165]:

(1) Center matrices $\mathbf{X}$ and $\mathbf{Y}$ by their means and scale them to unit variance.

(2) Set $\mathbf{u}_h$ equal to a column of $\mathbf{Y}$ with maximum variance.

(3) Set the dimension index $h = 1$.

(4) Regress columns of $\mathbf{X}$ on $\mathbf{u}_h$:

$$\mathbf{w}_h = \frac{\mathbf{X}\mathbf{u}_h}{\mathbf{u}_h^T \mathbf{u}_h} .$$

(5) Normalize $\mathbf{w}_h$ to unit length:

$$\mathbf{w}_h = \frac{\mathbf{w}_h}{\sqrt{\mathbf{w}_h^T \mathbf{w}_h}} .$$

(6) Calculate the input scores $\mathbf{s}_h$:

$$\mathbf{s}_h = \frac{\mathbf{X}\mathbf{w}_h}{\mathbf{w}_h^T \mathbf{w}_h} .$$

(7) Fit the nonlinear inner relation by least squares:

$$\mathbf{u}_h = c_{0h} + c_{1h}\mathbf{s}_h + c_{2h}\mathbf{s}_h^2 + \varepsilon_h ,$$

and calculate the estimated $\mathbf{u}_h$ ($\hat{\mathbf{u}}_h$):

$$\hat{\mathbf{u}}_h = c_{0h} + c_{1h}\mathbf{s}_h + c_{2h}\mathbf{s}_h^2 .$$

(8) Obtain the output loadings $\mathbf{q}_h$:

$$\mathbf{q}_h = \frac{\mathbf{Y}^T \hat{\mathbf{u}}_h}{\hat{\mathbf{u}}_h^T \hat{\mathbf{u}}_h} .$$

(9) Normalize $\mathbf{q}_h$to unit length:

$$\mathbf{q}_h = \frac{\mathbf{q}_h}{\sqrt{\mathbf{q}_h^T \mathbf{q}_h}} .$$

(10) Calculate new $\mathbf{u}_h$values:

$$\mathbf{u}_h = \frac{\mathbf{Y}\mathbf{q}_h}{\mathbf{q}_h^T \mathbf{q}_h} ,$$

then recalculate the coefficients in the inner relation c_{0h}, c_{1h}, and c_{2h} by least squares.

(11) Calculate correction to $\mathbf{w}$ as follows:

(a) Construct matrix $\mathbf{Z}_h$ with its first n_x columns equal to $(c_{1h} + 2c_{2h}\mathbf{s}_h) \otimes \mathbf{x}_k$ and its last three columns equal to 1, $\mathbf{s}_h$, and $\mathbf{s}_h^2$, where $\otimes$ means element by element multiplication and $k = 1, \ldots, n_x$.

(b) Obtain $\mathbf{v}_h$:

$$\mathbf{v}_h = \frac{\mathbf{Z}_h \mathbf{u}_h}{\mathbf{u}_h^T \mathbf{u}_h} .$$

(c) Normalize $\mathbf{v}_h$ to unit length:

$$\mathbf{v}_h = \frac{\mathbf{v}_h}{\sqrt{\mathbf{v}_h^T \mathbf{v}_h}} .$$

(d) Obtain $\mathbf{d}_h$:

$$\mathbf{d}_h = \frac{\mathbf{Z}_h \mathbf{v}_h}{\mathbf{v}_h^T \mathbf{v}_h} .$$

(e) Obtain b_h:

$$b_h = \frac{\mathbf{d}_h^T \mathbf{u}_h}{\mathbf{d}_h^T \mathbf{d}_h} .$$

(f) Update $\mathbf{w}_h$ as follows:

$$\mathbf{w}_h = \mathbf{w}_h + \boldsymbol{\Delta}\mathbf{w}_h \,,$$

where

$$\boldsymbol{\Delta}\mathbf{w}_h \left(j\right) = b\mathbf{v}_h \left(j\right) \;\text{ for}\, j = 1, \ldots, n_x \,.$$

(12) Calculate $\mathbf{s}_h$ according to step (6)

(13) Check convergence on $\mathbf{u}_h$. If the norm of the relative change in $\mathbf{u}_h$ with respect to the previous iteration is less than a specified tolerance (i.e. 10^{-10} is used in this study), proceed to step (14). Otherwise, repeat from step (7).

(14) Using the latest $\mathbf{s}_h$, calculate the final values of $\hat{\mathbf{u}}_h$, $\mathbf{q}_h$, $\mathbf{u}_h$ according to steps (7) to (10).

(15) Calculate the $\mathbf{X}$ loadings:

$$\mathbf{o}_h = \frac{\mathbf{X}\mathbf{s}_h}{\mathbf{s}_h^T \mathbf{s}_h} \,.$$

(16) Calculate the residuals of $\mathbf{X}$ and $\mathbf{Y}$:

$$\begin{aligned} \mathbf{E} &= \mathbf{X} - \mathbf{s}_h \mathbf{o}_h^T \,, \\ \mathbf{F} &= \mathbf{Y} - \mathbf{u}_h \mathbf{q}_h^T \,. \end{aligned}$$

If the current dimension h is less than the specified latent variables to be used (n_{latent}), set $h = h + 1$, use residuals $\mathbf{E}$ and $\mathbf{F}$ as $\mathbf{X}$ and $\mathbf{Y}$, respectively and repeat from step (4).

References

[1] ABDULLE, A. Fourth order Chebyshev methods with recurrence relation. *SIAM J. Sci. Comput. 23* (2002), 2041 – 2054.

[2] AHN, H. S., CHEN, Y. Q., AND MOORE, K. L. Iterative learning control: brief survey and categorization. *IEEE Transactions on Systems, Man and Cybernetics - Part C: Applications and Reviews 37*, 6 (2007), 1099 – 1121.

[3] BARD, Y. *Nonlinear parameter estimation*. New York: Academic Press, 1974.

[4] BATES, D. M., AND WATTS, D. G. *Nonlinear regression analysis and its applications*. New York: Wiley, 1988.

[5] BECK, J. V., AND ARNOLD, K. J. *Parameter estimation in engineering and science*. New York: Wiley, 1977.

[6] BLAGDEN, N., AND DAVEY, R. Polymorphs take shape. *Chem. Brit. 35* (1999), 44 – 47.

[7] BOIS, F. Y., FAHMY, T., BLOCK, J.-C., AND GATEL, D. Dynamic modeling of bacteria in a pilot drinking-water distribution system. *Water Res. 31*, 12 (1997), 3146 – 3456.

[8] BOTSARIS, G. D. *Secondary nucleation–a review*. Plenum Press, New York, 1976.

[9] BOX, G. E. P., AND DRAPER, N. R. The Bayesian estimation of common parameters from several responses. *Biometrika 52* (1965), 355 – 365.

[10] BOX, G. E. P., AND TIAO, G. C. *Bayesian inference in statistical analysis*. Reading, Mass: Addison-Wesley Pib. Co., 1973.

[11] BRAAK, C. J. A Markov Chain Monte Carlo version of the genetic algorithm differential evolution: easy Bayesian computing for real parameter spaces. *Stat. Comput. 16* (2006), 239 – 249.

[12] BRAATZ, R. D. Advanced control of crystallization processes. *Ann. Rev. Control 26* (2002), 87 – 99.

[13] BRAATZ, R. D., FUJIWARA, M., WUBBEN, T., AND RUSLI, E. Crystallization: particle size control. In *Encyclopedia of Pharmaceutical Technology*, J. Swarbrick, Ed., 3rd ed. New York: Marcel Dekker, 2006 (invited).

[14] BRETTHORST, G. L. An introduction to parameter estimation using Bayesian probability theory. In *Maximum Entropy and Bayesian Methods* (1990), P. F. Fougere, Ed., Dordrecht: Kluwer Academic Publishers.

[15] BRIESEN, H. Hierarchical characterization of aggregates for Monte Carlo simulations. *AIChE J. 52*, 7 (2006), 2436 – 2446.

[16] BRITTAIN, H. G. The impact of polymorphism on drug development: a regulatory viewpoint. *Amer. Pharm. Rev. 3* (2000), 67 – 70.

[17] BUDZ, J., JONES, A. G., AND MULLIN, J. W. On the shape-size dependence of potassium sulfate crystals. *Ind. Eng. Chem. Res. 26* (1987), 820 – 824.

[18] CAILLET, A., SHEIBAT-OTHMAN, N., AND FEVOTTE, G. Crystallization of monohydrate citric acid. 2. Modeling through population balance equations. *Cryst. Growth Des. 7* (2007), 2088 – 2095.

[19] CAMACHO, E. F., AND BORDONS, C. *Model predictive control*. London: Springer-Verlog, 1999.

[20] CANNING, T. F., AND RANDOLPH, A. D. Some aspects of crystallization theory: systems that violate McCabe's delta L law. *AIChE J. 13*, 1 (1967), 5 – 10.

[21] CARDEW, P. T., AND DAVEY, R. J. The kinetics of solvent-mediated phase transformations. *Proc. R. Soc. Lond. A 398*, 1815 (1985), 415 – 428.

[22] CARLIN, B. P., AND LOUIS, T. A. *Bayes and empirical Bayes methods for data analysis*. Boca Raton: Chapman & Hall/CRC, 2000.

[23] CHEN, W.-S., BAKSHI, B. R., GOEL, P. K., AND UNGARALA, S. Bayesian estimation via sequential Monte Carlo sampling: unconstrained nonlinear dynamic systems. *Ind. Eng. Chem. Res. 43* (2004), 4012 – 4025.

[24] CHIN, I., QIN, S. J., LEE, K. S., AND CHO, M. A two-stage iterative learning control technique combined with real-time feedback for independent disturbance rejection. *Automatica 40* (2004), 1913 – 1922.

[25] CHIN, I. S., LEE, K. S., AND LEE, J. H. A technique for integrated quality control, profile control, and constraint handling for batch processes. *Ind. Eng. Chem. Res. 39*, 3 (2000), 693 – 705.

[26] CLARKE-PRINGLE, T. L., AND MACGREGOR, J. F. Optimization of molecular-weight distribution using batch-to-batch adjustments. *Ind. Eng. Chem. Res. 37* (1998), 3660 – 3669.

[27] CLONTZ, N. A., AND MCCABE, W. L. Contact nucleation of magnesium sulfate heptahydrate. *AIChE Symp. Ser. 67*, 110 (1971), 6 – 17.

[28] COLEMAN, M. C., AND BLOCK, D. E. Bayesian parameter estimation with informative priors for nonlinear systems. *AIChE J. 52*, 2 (2006), 651 – 667.

[29] DAVEY, R., AND GARSIDE, J. *From molecules to crystallizers : an introduction to crystallization*. Oxford University Press, 2000.

[30] DE ANDA, J. C., WANG, X. Z., LAI, X., AND ROBERTS, K. J. Classifying organic crystals via in-process image analysis and the use of monitoring

charts to follow polymorphic and morphological changes. *J. Process Contr. 15*, 7 (2005), 785 – 797.

[31] DE ANDA, J. C., WANG, X. Z., AND ROBERTS, K. J. Multi-scale segmentation image analysis for the in-process monitoring of particle shape with batch crystallisers. *Chem. Eng. Sci. 60* (2005), 1053 – 1065.

[32] DE KEYSER, R., AND DONALD III, J. Application of the NEPSAC nonlinear predictive control strategy to a semiconductor reactor. *Lecture Notes in Control and Inform. Sci. 358* (2007), 407 – 417.

[33] DE KEYSER, R. M. C. Model based predictive control for linear systems. In *UNESCO Encyclopaedia of Life Support Systems*. Oxford: Eolss Publishers Co Ltd, 2003.

[34] DE KEYSER, R. M. C., AND CAUWENBERGHE, A. R. V. Extended prediction self-adaptive control. In *IFAC Symp. on Identification and System Parameter Estimate* (1985), pp. 1255 – 1260.

[35] DE PRADA, C., SARABIA, D., CRISTEA, S., AND MAZAEDA, R. Plant-wide control of a hybrid process. *Int. J. Adapt. Control 22* (2008), 124 – 141.

[36] DIEHL, M., BOCK, H. G., AND KOSTINA, E. An approximation technique for robust nonlinear optimization. *Math. Program., Ser. B 107* (2006), 213 – 230.

[37] DOYLE III, F. J., HARRISON, C. A., AND CROWLEY, T. J. Hybrid model-based approach to batch-to-batch control of particle size distribution in emulsion polymerization. *Comput. Chem. Eng. 27* (2003), 1153 – 1163.

[38] DURAN, M. A., AND WHITE, B. S. Bayesian estimation applied to effective heat transfer coefficients in a packed bed. *Chem. Eng. Sci. 50*, 3 (1995), 495 – 510.

[39] FATEMI, E., JEROME, J., AND OSHER, S. Solution of the hydrodynamic device model using high order non-oscillatory shock capturing algorithms. *IEEE Trans. Comput. Aid. Des. 10* (1991), 232 – 244.

[40] FEVOTTE, G., ALEXANDRE, C., AND NIDA, S. O. A population balance model of the solution-mediated phase transition of citric acid. *AIChE J. 53* (2007), 2578 – 2589.

[41] FUJIWARA, M., CHOW, P. S., MA, D. L., AND BRAATZ, R. D. Paracetamol crystallization using laser backscattering and ATR-FTIR spectroscopy: metastability, agglomeration and control. *Cryst. Growth Des. 2*, 5 (2002), 363 – 370.

[42] FUJIWARA, M., NAGY, Z. K., CHEW, J. W., AND BRAATZ, R. D. First-principles and direct design approaches for the control of pharmaceutical crystallization. *J. Process Contr. 15* (2005), 493 – 504.

[43] GARCIA, C. E., AND MORSHEDI, A. M. Quadratic programming solution of dynamic matrix control (QDMC). *Chem. Eng. Comm. 46* (1986), 73 – 87.

[44] GARCIA, C. E., PRETT, D. M., AND MORARI, M. Model predictive control: theory and practice - a survey. *Automatica 25*, 3 (1989), 1753 – 1758.

[45] GARSIDE, J., PHILLIPS, V. R., AND SHAH, M. B. On size-dependent crystal growth. *Ind. Eng. Chem. Fundam. 15*, 3 (1976), 230 – 233.

[46] GARSIDE, J., AND SHAH, M. B. Crystallization kinetics from MSMPR crystallizers. *Ind. Eng. Chem. Process Des. Dev. 19* (1980), 509 – 514.

[47] GELMAN, A., CARLIN, J. B., STERN, H. S., AND RUBIN, D. B. *Bayesian data analysis*. New York: Chapman & Hall/CRC, 2004.

[48] GIROLAMI, M. W., AND ROUSSEAU, R. W. Size-dependent crystal growth - a manifestation of growth rate dispersion in the potassium alum-water system. *AIChE J. 31*, 11 (1985), 1821 – 1828.

[49] GREENOUGH, J., AND RIDER, W. A quantitative comparision of numerical methods for the compressible Euler equations: fifth-order WENO and piecewise-linear Godunov. *J. Comput. Phys. 196* (2004), 259 – 281.

[50] GRON, H., BORISSOVA, A., AND ROBERTS, K. J. In-process ATR-FTIR spectroscopy for closed-loop supersaturation control of a batch crystallizer producing monosodium glutamate crystals of defined size. *Ind. Eng. Chem. Res. 42* (2003), 198 – 206.

[51] GROOTSCHOLTEN, P. A. M., DE LEER, B. G. M., DE JONG, E. J., AND ASSELBERGS, C. Factors affecting secondary nucleation rate of sodium chloride in an evaporative crystallizer. *AIChE J. 28* (1982), 728 – 737.

[52] GUNAWAN, R., FUSMAN, I., AND BRAATZ, R. D. High resolution algorithms for multidimensional population balance equations. *AIChE J. 50*, 11 (2004), 2738 – 2748.

[53] GUNAWAN, R., JUNG, M. Y., SEEBAUER, E. G., AND BRAATZ, R. D. Maximum a posteriori estimation of transient enhanced diffusion energetics. *AIChE J. 49*, 8 (2003), 2114 – 2122.

[54] GUNAWAN, R., MA, D. L., FUJIWARA, M., AND BRAATZ, R. D. Identification of kinetic parameters in a multidimensional crystallization process. *Int. J. Mod. Phys. B 16* (2002), 367 – 374.

[55] HAARIO, H., SAKSMAN, E., AND TAMMINEN, J. An adaptive Metropolis algorithm. *Bernoulli 7* (2001), 223 – 242.

[56] HARTEN, A. High resolution schemes for hyperbolic conservation laws. *J. Comput. Phys. 49* (1983), 357 – 393.

[57] HARTEN, A., ENGQUIST, B., OSHER, S., AND CHAKRAVARTHY, S. Uniformly high order essentially non-oscillatory schemes, III. *J. Comput. Phys. 71* (1987), 231 – 303.

[58] HASELTINE, E. L., PATIENCE, D. B., AND RAWLINGS, J. B. On the stochastic simulation of particulate systems. *Chem. Eng. Sci. 60* (2005), 2627 – 2641.

[59] HEATH, A. R., FAWELL, P. D., BAHRI, P. A., AND SWIFT, J. D. Estimating average particle size by focused beam reflectance measurement (FBRM). *Part. Part. Syst. Charact. 19* (2002), 84 – 95.

[60] HELT, J. E., AND LARSON, M. A. Effect of temperature on the crystallization of potassium nitrate by direct measurement of supersaturation. *AIChE J. 23*, 6 (1977), 822 – 830.

[61] HENRICK, A. K., ASLAM, T. D., AND POWERS, J. M. Mapped weighted essentially non-oscillatory schemes: achieving optimal order near critical points. *J. Comput. Phys. 207* (2005), 542 – 567.

[62] HENSON, M. A. Nonlinear model predictive control: current status and future directions. *Comput. Chem. Eng. 23* (1998), 187 – 202.

[63] HENSON, M. A., AND SEBORG, D. E., Eds. *Nonlinear process control*, 1st ed. Piscataway, NJ: Prentice Hall, 1996.

[64] HERMANTO, M. W., CHIU, M.-S., WOO, X.-Y., AND BRAATZ, R. D. Robust optimal control of polymorphic transformation in batch crystallization. *AIChE J. 53*, 10 (2007), 2643 – 2650.

[65] HOUNSLOW, M. J., RYALL, R. L., AND MARSHALL, V. R. A discretized population balance for nucleation, growth, and aggregation. *AIChE J. 34* (1988), 1821 – 1832.

[66] HU, Q., ROHANI, S., AND JUTAN, A. Modelling and optimization of seeded batch crystallizers. *Comput. Chem. Eng. 29* (2005), 911 – 918.

[67] HU, Q., ROHANI, S., WANG, D. X., AND JUTAN, A. Nonlinear kinetic parameter estimation for batch cooling seeded crystallization. *AIChE J. 50*, 8 (2004), 1786 – 1794.

[68] HUKKANEN, E. J., AND BRAATZ, R. D. Measurement of particle size distribution in suspension polymerization using in situ laser backscattering. *Sens. Actuators B 96* (2003), 451 – 459.

[69] HULBURT, H. M., AND KATZ, S. Some problems in particle technology: a statistical mechanical formulation. *Chem. Eng. Sci. 19* (1964), 555 – 574.

[70] IONESCU, C., AND DE KEYSER, R. M. C. EPSAC predictive control of blood glucose level in type I diabetic patients. In *Proceedings of the 44th Conference on Decision and Control, and the European Control Conference* (2005), pp. 4845 – 4850.

[71] JANSE, A. H., AND DE JONG, E. J. *The occurence of growth dispersion and its consequences*. New York: Plenum Press, 1976.

[72] JIANG, G.-S., AND SHU, C.-W. Efficient implementation of weighted ENO schemes. *J. Comput. Phys. 126* (1996), 202 – 228.

[73] JOHNSON, R. T., ROUSSEAU, R. W., AND MCCABE, W. L. Factors affecting contact nucleation. *AIChE Symp. Ser. 68*, 121 (1972), 31 – 41.

[74] JULIER, S. J. The scaled unscented transformation. In *Proceedings of the American Control Conference* (2002), vol. 6, pp. 4555 – 4559.

[75] JULIER, S. J. The spherical simplex unscented transformation. In *Proceedings of the American Control Conference* (2003), vol. 3, pp. 2430 – 2434.

[76] JULIER, S. J., AND UHLMANN, J. K. Reduced sigma point filters for the propagation of means and covariances through nonlinear transformation. In *Proceedings of the American Control Conference* (2002), vol. 2, pp. 887 – 892.

[77] JULIER, S. J., AND UHLMANN, J. K. Unscented filtering and nonlinear estimation. *Proceedings of the IEEE 92*, 3 (2004), 401 – 422.

[78] JULIER, S. J., UHLRNANN, J. K., AND DURRANT-WHYTE, H. F. A new approach for filtering nonlinear systems. In *Proceedings of the American Control Conference* (1995), pp. 1628 – 1632.

[79] KALBASENKA, A. N., SPIERINGS, L. C. P., HUESMAN, A. E. M., AND KRAMER, H. J. M. Application of seeding as a process actuator in a model

predictive control framework for fed-batch crystallization of ammonium sulphate. *Part. Part. Syst. Charact. 24* (2007), 40 – 48.

[80] KEE, N., WOO, X. Y., GOH, L., CHEN, K., HE, G., BHAMIDI, V., RUSLI, E., NAGY, Z. K., KENIS, P. J. A., ZUKOSKI, C. F., TAN, R. B. H., AND BRAATZ, R. D. Design of crystallization processes from laboratory R&D to the manufacturing scale. *Pharmaceuticals Review* (2007 (to be published)).

[81] KITAMURA, M., AND ISHIZU, T. Growth kinetics and morphological change of polymorphs of L-glutamic acid. *J. Cryst. Growth 209* (2000), 138 – 145.

[82] KUMAR, S., AND RAMKRISHNA, D. On the solution of population balance equations by discretization-I. A fixed pivot technique. *Chem. Eng. Sci. 51* (1996), 1311 – 1332.

[83] KUMAR, S., AND RAMKRISHNA, D. On the solution of population balance equations by discretization-III. Nucleation, growth and aggregation of particles. *Chem. Eng. Sci. 52* (1997), 4659 – 4679.

[84] LAMPINEN, J. A constraint handling approach for the differential evolution algorithm. In *Proceedings of the 2002 Congress on Evolutionary Computation* (2002), vol. 2, pp. 1468 – 1473.

[85] LANG, L., CHEN, W.-S., BAKSHI, B. R., GOEL, P. K., AND UNGARALA, S. Bayesian estimation via sequential Monte Carlo sampling - constrained dynamic systems. *Automatica 43* (2007), 1615 – 1622.

[86] LARSEN, P. A., PATIENCE, D. B., AND RAWLINGS, J. B. Industrial crystallization process control. *IEEE Contr. Syst. Mag. 26*, 4 (2006), 70 – 80.

[87] LASKEY, K. B., AND MYERS, J. W. Population Markov Chain Monte Carlo. *Machine Learning 50* (2003), 175 – 196.

[88] LEE, J. H., AND LEE, K. S. Iterative learning control applied to batch processes: an overview. *Control Eng. Pract. 15* (2007), 1306 – 1318.

[89] LEE, J. H., LEE, K. S., AND KIM, W. C. Model-based iterative learning control with a quadratic criterion for time-varying linear systems. *Automatica 36*, 5 (2000), 641 – 657.

[90] LEE, K., LEE, J. H., YANG, D. R., AND MAHONEY, A. W. Integrated run-to-run and on-line model-based control ofparticle size distribution for a semi-batch precipitation reactor. *Comput. Chem. Eng. 26* (2002), 1117 – 1131.

[91] LEE, K. S., CHIN, I. S., LEE, H. J., AND LEE, J. H. Model predictive control technique combined with iterative learning for batch processes. *AIChE J. 45*, 10 (1999), 2175 – 2187.

[92] LEVEQUE, R. J. *Numerical methods for conservation laws*. Basel, Germany: Birkhauser Verlag, 1992.

[93] LEVEQUE, R. J. Wave propagation algorithms for multidimensional hyperbolic systems. *J. Comput. Phys. 131* (1997), 327 – 353.

[94] LI, C., ZHANG, J., AND WANG, G. Batch-to-batch optimal control of batch processes based on recursively updated nonlinear partial least squares models. *Chem. Eng. Comm. 194*, 3 (2007), 261 – 279.

[95] LIANG, F. Dynamically weighted importance sampling in Monte Carlo computation. *J. Am. Stat. Assoc. 97* (2002), 807 – 821.

[96] LIANG, F. M., AND WONG, W. H. Real-parameter evolutionary Monte Carlo with applications to Bayesian mixture models. *J. Am. Stat. Assoc. 96* (2001), 653 – 666.

[97] LIOTTA, V., AND SABESAN, V. Monitoring and feedback control of supersaturation using ATR-FTIR to produce an active pharmaceutical ingredient of a desired crystal size. *Org. Process Res. Dev. 8* (2004), 488 – 494.

[98] LIU, J. S. *Monte Carlo strategies in scientific computing*. New York: Springer, 2001.

[99] LIU, X.-D., OSHER, S., AND CHAN, T. Weighted essentially non-oscillatory schemes. *J. Comput. Phys. 115* (1994), 200 – 212.

[100] MA, C. Y., WANG, X. Z., AND ROBERTS, K. J. Multi-dimensional population balance modeling of the growth of rod-like L-glutamic acid crystals using growth rates estimated from in-process imaging. *Adv. Powder Technol. 18* (2007), 707 – 723.

[101] MA, C. Y., WANG, X. Z., AND ROBERTS, K. J. Morphological population balance for modeling crystal growth in face directions. *AIChE J. 54*, 1 (2008), 209 – 222.

[102] MA, D. L., AND BRAATZ, R. D. Worst-case analysis of finite-time control policies. *IEEE Trans. Contr. Syst. Technol. 9*, 5 (2001), 766 – 774.

[103] MA, D. L., CHUNG, S. H., AND BRAATZ, R. D. Worst-case performance analysis of optimal batch control trajectories. *AIChE J. 45*, 7 (1999), 1469 – 1476.

[104] MA, D. L., TAFTI, D. K., AND BRAATZ, R. D. High-resolution simulation of multidimensional crystal growth. *Ind. Eng. Chem. Res. 41* (2002), 6217 – 6223.

[105] MATTHEWS, H. B., AND RAWLINGS, J. B. Batch crystallization of a photochemical: modeling, control, and filtration. *AIChE J. 44*, 5 (1998), 1119 – 1127.

[106] MENDES, P., AND KELL, D. B. Non-linear optimization of biochemical pathways: applications to metabolic engineering and parameter estimation. *Bioinformatics 14*, 10 (1998), 869 – 883.

[107] MERSMANN, A. *Crystallization technology handbook*, 2nd ed. Florida, USA: CRC Press, 2001.

[108] METROPOLIS, N., ROSENBLUTH, A. W., ROSENBLUTH, M. N., AND TELLER, A. H. Equation of state calculations by fast computing machines. *J. Chem. Phys. 21*, 6 (1953), 1087 – 1092.

[109] MILLER, S. M., AND RAWLINGS, J. B. Model identification and control strategies for batch cooling crystallizers. *AIChE J. 40*, 8 (1994), 1312 – 1327.

[110] MORARI, M., AND LEE, J. H. Model predictive control: past, present and future. *Comput. Chem. Eng. 23* (1999), 667 – 682.

[111] MORRIS, K. R., GRIESSER, U. J., ECKHARDT, C. J., AND STOWELL, J. G. Theoretical approaches to physical transformations of active pharmaceutical ingredients during manufacturing processes. *J. Cryst. Growth 48* (2001), 91 – 114.

[112] MULLIN, J. W. *Crystallization*. London: Butterworths, 1961.

[113] NAGY, Z. K., AND BRAATZ, R. D. Robust nonlinear model predictive control of batch processes. *AIChE J. 49*, 7 (2003), 1776 – 1786.

[114] NAGY, Z. K., AND BRAATZ, R. D. Worst-case and distributional robustness analysis of finite-time control trajectories for nonlinear distributed parameter systems. *IEEE Trans. Contr. Syst. Technol. 11*, 5 (2003), 694 – 704.

[115] ONO, T., KRAMER, H. J. M., TER HORST, J. H., AND JANSENS, P. J. Process modeling of the polymorphic transformation of L-glutamic acid. *Cryst. Growth Des. 4*, 6 (2004), 1161 – 1167.

[116] OSHER, S., AND CHAKRAVARTHY, S. High resolution schemes and the entropy condition. *SIAM J. Numer. Anal. 21* (1984), 955 – 984.

[117] POUILLOT, R., ALBERT, I., CORNU, M., AND DENIS, J.-B. Estimation of uncertainty and variability in bacterial growth using Bayesian inference. Application to listeria monocytogenes. *Int. J. Food Microbiol. 81* (2003), 87 – 104.

[118] PUEL, F., FEVOTTE, G., AND KLEIN, J. P. Simulation and analysis of industrial crystallization processes through multidimensional population balance equations. Part 1: a resolution algorithm based on the method of classes. *Chem. Eng. Sci. 58*, 6 (2003), 3715 – 3727.

[119] QAMAR, S., ASHFAQA, A., WARNECKEA, G., ANGELOVC, I., ELSNERC, M., AND SEIDEL-MORGENSTERN, A. Adaptive high-resolution schemes for multidimensional population balances in crystallization processes. *Comput. Chem. Eng. 31*, 10 (2007), 1296 – 1311.

[120] QAMAR, S., ELSNER, M. P., ANGELOV, I., WARNECKE, G., AND SEIDEL-MORGENSTERN, A. A comparative study of high resolution schemes for solving population balances in crystallization. *Comput. Chem. Eng. 30*, 6 – 7 (2006), 1119 – 1131.

[121] QAMAR, S., AND WARNECKEA, G. Numerical solution of population balance equations for nucleation, growth and aggregation processes. *Comput. Chem. Eng. 31* (2007), 1576 – 1589.

[122] QIN, S. J., AND BADGWELL, T. A. An overview of industrial model predictive control technology. In *Chemical Process Control - AIChE Symposium Series*, J. Kantor, C. E. Garcia, and B. Carnahan, Eds. New York: AIChE, 1997, pp. 232 – 256.

[123] RAMKRISHNA, D. The status of population balances. *Rev. Chem. Eng. 3* (1985), 49 – 95.

[124] RANDOLPH, A. D., AND LARSON, M. A. Transient and steady state size distributions in continuous mixed suspension crystallizers. *AIChE J. 8*, 5 (1962), 639 – 645.

[125] RANDOLPH, A. D., AND LARSON, M. A. *Theory of particulate processes : analysis and techniques of continuous crystallization*. San Diego: Academic Press, 1988.

[126] RANDOLPH, A. D., AND WHITE, E. T. Modeling size dispersion in the prediction of crystal-size distribution. *Chem. Eng. Sci. 32* (1977), 1067 – 1076.

[127] RAWLINGS, J. B. Tutorial overview of model predictive control. *IEEE Contr. Syst. Mag. 20*, 3 (2000), 38 – 52.

[128] RAWLINGS, J. B., MILLER, S. M., AND WITKOWSKI, W. R. Model identification and control of solution crystallization processes: a review. *Ind. Eng. Chem. Res. 32* (1993), 1275 – 1296.

[129] RODRIGUES-HORNEDO, N., AND MURPHY, D. Significance of controlling crystallization mechanisms and kinetics in pharmaceutical systems. *J. Pharm. Sci. 88* (1999), 651 – 660.

[130] ROELANDS, C. P., JIANG, S. F., KITAMURA, M., TER HORST, J. H., KRAMER, H. J., AND JANSENS, P. J. Antisolvent crystallization of the polymorphs of L-histidine as a function of supersaturation ratio and of solvent composition. *Cryst. Growth Des. 6*, 4 (2006), 955 – 963.

[131] ROHANI, S., HAERI, M., AND WOOD, H. C. Modeling and control of a continuous crystallization process - part 2. Model predictive control. *Comput. Chem. Eng. 23* (1999), 279 – 286.

[132] ROHANI, S., HORNE, S., AND MURTHY, K. Control of product quality in batch crystallization of pharmaceuticals and fine chemicals. Part 1: design of

the crystallization process and the effect of solvent. *Org. Process Res. Dev. 9* (2005), 858 – 872.

[133] ROHANI, S., HORNE, S., AND MURTHY, K. Control of product quality in batch crystallization of pharmaceuticals and fine chemicals. Part 2: external control. *Org. Process Res. Dev. 9* (2005), 873 – 883.

[134] RUEDA, A., CRISTEA, S., PRADA, C. D., AND DE KEYSER, R. M. C. Non-linear predictive control for a distillation column. In *Proceedings of the 44th IEEE conference on Decision and Control, and the European Control Conference* (2005), pp. 5156 – 5161.

[135] RUF, A., WORLITSCHEK, J., AND MAZZOTTI, M. Modeling and experimental analysis of PSD measurements through FBRM. *Part. Part. Syst. Charact. 17* (2000), 167 – 179.

[136] SAKAI, H., HOSOGAI, H., KAWAKITA, T., ONUMA, K., AND TSUKAMOTO, K. Transformation of α-glycine to γ-glycine. *J. Cryst. Growth 116* (1992), 421 – 426.

[137] SARANTEAS, K., BAKALE, R., HONG, Y. P., LUONG, H., FOROUGHI, R., AND WALD, S. Process design and scale-up elements for solvent mediated polymorphic controlled tecastemizole crystallization. *Org. Process Res. Dev. 9* (2005), 911 – 922.

[138] SCHIESSER, W. E. *The numerical method of lines - integration of partial differential equations.* New York: Academic Press, 1991.

[139] SCHOLL, J., BONALUMI, D., VICUM, L., AND MAZZOTTI, M. In situ monitoring and modeling of the solvent-mediated polymorphic transformation of L-glutamic acid. *Cryst. Growth Des. 6*, 4 (2006), 881 – 891.

[140] SCOKAERT, P. O. M., AND RAWLINGS, J. B. Feasibility issues in linear model predictive control. *AIChE J. 45*, 8 (1999), 1649 – 1659.

[141] SERNA, S., AND MARQUINA, A. Power ENO methods: a fifth-order accurate weighted power ENO method. *J. Comput. Phys. 194* (2004), 632 – 658.

[142] SHEKUNOV, B. Y., AND YORK, P. Crystallization processes in pharmaceutical technology and drug delivery design. *J. Cryst. Growth 211* (2000), 122 – 136.

[143] SHU, C.-W. Efficient implementation of essentially non-oscillatory shock-capturing schemes. *J. Comput. Phys. 77* (1988), 439 – 471.

[144] SHU, C.-W. Efficient implementation of essentially non-oscillatory shock-capturing schemes, II. *J. Comput. Phys. 83* (1989), 32 – 78.

[145] SHU, C. W. Numerical experiments on the accuracy of ENO and modified ENO schemes. *J. Sci. Comput. 5* (1990), 127 – 149.

[146] SIMMONS, M., LANGSTON, P., AND BURBIDGE, A. Particle and droplet size analysis from chord distributions. *Powder Technol. 102* (1999), 75–83.

[147] SINGH, P. N., AND RAMKRISHNA, D. Solution of population balance equations by MWR. *Comput. Chem. Eng. 1* (1977), 23 – 31.

[148] SOTOWA, K.-I., NAITO, K., KANO, M., HASEBE, S., AND HASHIMOTO, I. Application of the method of characteristics to crystallizer simulation and comparison with finite difference for controller performance evaluation. *J. Process Contr. 10* (2000), 203 – 208.

[149] SRINIVASAN, B., BONVIN, D., VISSER, E., AND PALANKI, S. Dynamic optimization of batch processes - II. Role of measurements in handling uncertainty. *Comput. Chem. Eng. 27* (2003), 27 – 44.

[150] STARBUCK, C., SPARTALIS, A., WAI, L., WANG, J., FERNANDEZ, P., LINDEMANN, C. M., ZHOU, G. X., AND GE, Z. H. Process optimization of a complex pharmaceutical polymorphic system via in situ Raman spectroscopy. *Cryst. Growth Des. 2* (2002), 515 – 522.

[151] STORN, R., AND PRICE, K. Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces. *J. Global Optim. 11* (1997), 341 – 359.

[152] STRICKLAND-CONSTABLE, R. F. The breeding of crystal nuclei - a review of the subject. *AIChE Symp. Ser. 68,* 121 (1972), 1 – 7.

[153] SWEBY, P. K. High resolution schemes using flux limiters for hyperbolic conservation laws. *SIAM J. Numer. Anal. 21*, 5 (1984), 995 – 1011.

[154] TADAYYON, A., AND ROHANI, S. Determination of particle size distribution by Par-Tec 100: modeling and experimental results. *Part. Part. Syst. Charact. 15* (1998), 127 – 135.

[155] TADAYYON, A., AND ROHANI, S. Extended Kalman filter-based nonlinear model predictive control of a continuous KCl-NaCl crystallizer. *Can. J. Chem. Eng 79* (2001), 255 – 262.

[156] TAMAS, L., NASCU, I., AND DE KEYSER, R. The NEPSAC nonlinear predictive controller in a real life experiment. In *International Conference on Intelligent Engineering Systems* (2007), pp. 229 – 234.

[157] TIERNEY, L. Markov chains for exploring posterior distributions. *Ann. Stat. 22*, 4 (1994), 1701 – 1728.

[158] TOGKALIDOU, T., BRAATZ, R. D., JOHNSON, B., DAVIDSON, O., AND ANDREWS, A. Experimental design and inferential modeling in pharmaceutical crystallization. *AIChE J. 47*, 1 (2001), 160 – 168.

[159] TOGKALIDOU, T., FUJIWARA, M., PATEL, S., AND BRAATZ, R. D. Solute concentration prediction using chemometrics and ATR-FTIR spectroscopy. *J. Cryst. Growth 231* (2001), 534 – 543.

[160] TOGKALIDOU, T., TUNG, H. H., SUN, Y., ANDREWS, A., AND BRAATZ, R. D. Solution concentration prediction for pharmaceutical crystallization processes using robust chemometrics and ATR FTIR spectroscopy. *Org. Process Res. Dev. 6*, 6 (2002), 317 – 322.

[161] TOGKALIDOU, T., TUNG, H. H., SUN, Y., ANDREWS, A. T., AND BRAATZ, R. D. Parameter estimation and optimization of a loosely-bound aggregating pharmaceutical crystallization using in-situ infrared and laser backscattering measurements. *Ind. Eng. Chem. Res. 43* (2004), 6168 – 6181.

[162] VAN LEER, B. Towards the ultimate conservative difference scheme II. Monotonicity and conservation combined in a second order scheme. *J. Comput. Phys. 14* (1974), 361 – 370.

[163] WALTON, A. G. *Nucleation in liquids and solutions*. New York: Marcel Dekker, 1969.

[164] WHITE, E. T., AND WRIGHT, P. G. Magnitude of size dispersion effects in crystallization. *Chem. Eng. Prog. Symp. Ser. 67*, 110 (1971), 81 – 87.

[165] WOLD, S., KETTANEH-WOLD, N., AND SKAGERBERG, B. Nonlinear PLS modeling. *Chemom. Intell. Lab. Syst. 7* (1989), 53 – 65.

[166] WOO, X. Y., TAN, R. B. H., CHOW, P. S., AND BRAATZ, R. D. Simulation of mixing effects in antisolvent crystallization using a coupled CFD-PDF-PBE approach. *Cryst. Growth Des. 6*, 6 (2006), 1291 – 1303.

[167] WORLITSCHEK, J., AND MAZZOTTI, M. Model-based optimization of particle size distribution in batch-cooling crystallization of paracetamol. *Cryst. Growth Des. 4*, 5 (2004), 891 – 903.

[168] XIE, Y.-L., AND KALIVAS, J. H. Evaluation of principal component selection methods to form a global prediction model by principal component regression. *Anal. Chim. Acta 348* (1997), 19 – 27.

[169] XIONG, Z., AND ZHANG, J. Product quality trajectory tracking in batch processes using iterative learning control based on time-varying perturbation models. *Ind. Eng. Chem. Res. 42* (2003), 6802 – 6814.

[170] XIONG, Z., AND ZHANG, J. A batch-to-batch iterative optimal control strategy based on recurrent neural network models. *J. Process Contr. 15* (2005), 11 – 21.

[171] YU, L. Survival of the fittest polymorph: how fast nucleater can lose to fast grower. *CrystEngComm 9* (2007), 841 – 851.

[172] YU, L. X., LIONBERGER, R. A., RAW, A. S., D'COSTA, R., WU, H. Q., AND HUSSAIN, A. S. Applications of process analytical technology to crystallization processes. *Adv. Drug Delivery Rev. 56*, 3 (2004), 349 – 369.

[173] ZAFIRIOU, E., ADOMAITIS, R. A., AND GATTU, G. An approach to run-to-run control for rapid thermal processing. In *Proceedings of the American Control Conference* (1995), pp. 1286 – 1288.

[174] ZHANG, G. P., AND ROHANI, S. On-line optimal control of a seeded batch cooling crystallizer. *Chem. Eng. Sci. 58* (2003), 1887 – 1896.

[175] ZHOU, G. X., FUJIWARA, M., WOO, X. Y., RUSLI, E., TUNG, H. H., STARBUCK, C., DAVIDSON, O., GE, Z., AND BRAATZ, R. D. Direct design of pharmaceutical antisolvent crystallization through concentration control. *Cryst. Growth Des. 6* (2006), 892 – 898.

[176] ZUMSTEIN, R. C., AND ROUSSEAU, R. W. Growth rate dispersion by initial growth rate distributions and growth rate fluctuations. *AIChE J. 33*, 1 (1987), 121 – 129.

Publications and Presentations

HERMANTO, M.W., BRAATZ, R.D., CHIU, M.-S. Robust optimal control of polymorphic transformation in batch crystallization. *AIChE J. 53*, 10 (2007), 2643 – 2650.

HERMANTO, M. W., KEE, N. C., BRAATZ, R. D., CHIU, M.-S., TAN, R. B. H. Robust Bayesian estimation of kinetics for the polymorphic transformation of L-glutamic acid crystals. *AIChE J. 54*, 12 (2008), 3248 – 3259.

HERMANTO, M.W., BRAATZ, R.D., CHIU, M.-S. High-order simulation of polymorphic crystallization using weighted essentially non-oscillatory methods. *AIChE J. 55*, 1 (2009), 122 – 131.

HERMANTO, M.W., BRAATZ, R.D., CHIU, M.-S. Nonlinear model predictive control for the polymorphic transformation of L-glutamic acid crystals. *AIChE J.* (2008), in press.

HERMANTO, M.W., BRAATZ, R.D., CHIU, M.-S. Nonlinear MPC technique combined with batch-to-batch control for polymorphic transformation in pharma-

ceutical crystallization, under preparation.

HERMANTO, M.W., BRAATZ, R.D., CHIU, M.-S. Run-to-run temperature control for polymorphic transformation in pharmaceutical crystallization with uncertainties. In *The 6th World Congress on Intelligent Control and Automation* (2006), Dalian, China, 21-23 June.

HERMANTO, M.W., BRAATZ, R.D., CHIU, M.-S. A comparative study of temperature control and concentration control for polymorphic transformation in pharmaceutical crystallization with uncertainties. In *The 6th Asian Control Conference* (2006), Bali, Indonesia, 18-21 July.

HERMANTO, M.W., BRAATZ, R.D., CHIU, M.-S. A run-to-run control strategy for polymorphic transformation in pharmaceutical crystallization. In *IEEE Conference on Control Applications, and Automation* (2006), Munich, Germany, 4-6 October.

HERMANTO, M.W., BRAATZ, R.D., CHIU, M.-S. Simulation and process modelling of polymorphic crystallization. In *Annual Graduate Student Symposium in Biological and Chemical Engineering* (2007), Singapore, 14 September.

HERMANTO, M.W., BRAATZ, R.D., CHIU, M.-S. Optimal control of polymorphic transformation in batch pharmaceutical crystallization. In *IEEE Conference on Control Applications* (2007), Singapore, 1-3 October.

HERMANTO, M.W., KEE N.C., BRAATZ, R.D., CHIU, M.-S., TAN R.B.H. Mod-

eling the polymorphic crystallization of L-glutamic acid by Bayesian parameter estimation. In *International Congress of Chemical and Process Engineering* (2008), Prague, Czech Republic, 24 - 28 August.

HERMANTO, M.W., BRAATZ, R.D., CHIU, M.-S. Modelling and simulation of the crystallization of L-glutamic acid polymorphs. In *AIChE Annual Meeting* (2008), Philadelphia, USA, 16-21 November.