Kinetic Parameter Estimation from Spectroscopic Data for a Multi-Stage Solid-Liquid Pharmaceutical Process[†]

Christina Schenk,[*,‡,¶] Lorenz T. Biegler,[¶] Lu Han,[*,§] and Jason Mustakis[§]

‡*BCAM - Basque Center for Applied Mathematics, Mazarredo 14, E48009 Bilbao, Basque Country - Spain*

¶*Carnegie Mellon University, Chemical Engineering Department, Pittsburgh, PA 15213 - USA*

§*Pfizer Inc., Groton, CT 06340 - USA*

E-mail: cschenk@bcamath.org; lu.han2@pfizer.com

Abstract

Laboratory and process measurements from spectroscopic instruments are ubiquitous in pharma processes, and directly using the data can pose a number of challenges for kinetic model building. Moreover, scaling up from laboratory to industrial level requires predictive models with accurate parameter values. This means that process identification does not only imply kinetic parameter estimation, but also the identification of the absorbing species and estimation of variances for both the data and parameters. A recently developed, open-source toolkit *KIPET*[1,2] addresses these topics and provides an alternative to standard parameter estimation packages, in particular for spectroscopic data problems. Moreover, batch processes commonly used in the

[†]Preprint submitted to Organic Process Research & Development

chemical and pharmaceutical industry involve multiple stages to carry out synthesis operations in a step by step manner, often dealing with heterogeneous mixtures, wide operating temperatures, and constant additions and removals of product and waste. For such cases novel modeling approaches are required, as the structure of the kinetic model may vary with time, with model switches that are state dependent. This study presents a new modeling approach and methodology that deals with these practical issues. In developing kinetic models, it approximates the solid dissolution process and deals with multiple stages with different reactor temperatures. Moreover, variances, parameters, concentration and absorbance profiles are estimated for the process stages using the approach presented by Chen et al.[3]. The application of these developed concepts results in realistic profiles as well as reliable kinetic parameter values. The outcomes of this work show that *KIPET* is a useful toolkit for dealing with pharmaceutical processes with capabilities for dealing with challenging kinetic modeling problems.

1 Introduction

For chemical and pharmaceutical industries, it is crucial to guarantee controllability, safety and scalability of manufacturing processes. Usually in the early development phase, the reactions are not completely understood, and the underlying reaction network, chemical species and kinetic parameters involved are unclear. Typical data collected during these experiments include spectroscopic data and high-performance or ultra-performance liquid chromatography data. For the pharma example presented in this paper, parameter estimation based on infrared (IR) spectroscopic data is considered. Several approaches are described in the literature that deal with spectroscopic data, such as spectroscopic multivariate curve resolution (MCR) techniques[4,5] connected through Beer-Lambert's law. MCR's use is widespread in industry for a variety of applications. Its goal is the decomposition of the data matrix into its bilinear components. Beer-Lambert's law describes the relationship

Acronyms

CI	Confidence Interval	NLP	Nonlinear Programming
DAE	Differential Algebraic Equation	PCA	Principal Component Analysis
IPOPT	Interior Point OPTimizer	PCD	Pure Component Decomposition
IR	Infrared		
KIPET	Kinetic Parameter Estimation Toolkit	Pyomo	Python optimization modeling objects
LoF	Lack of Fit		
MCR	Multivariate Curve Resolution	sIPOPT	Sensitivity Based on IPOPT
MCR-ALS	MCR-Alternating Least Squares	SVD	Singular Value Decomposition

of the absorbance profiles with the pure components, i.e.

$$\mathbf{D} = \mathbf{C}\mathbf{S}^{\mathbf{T}} + \mathbf{E}, \tag{1}$$

where $\mathbf{D}$ is the measured spectroscopic data given in matrix form of dimension $ntp \times nwp$ with time points t_i, $i = 1, \ldots, ntp$ and wavelengths λ_l, $l = 1, \ldots, nwp$. Moreover $\mathbf{C}$ represents the molar concentration profile in matrix form of dimension $ntp \times nc$ with species c_k, $k = 1, \ldots, nc$. The $nc \times nwp$-dimensional matrix $\mathbf{S}$ represents the absorbance profiles. The remaining $ntp \times nwp$-dimensional matrix $\mathbf{E}$ describes the measurement error. Lawton and Sylvestre[4],[5] aimed to estimate $\mathbf{C}$ and $\mathbf{S}$ simultaneously using a soft-modeling approach, where the addition of non-negativity constraints for $\mathbf{C}$ and $\mathbf{S}$ yields physical meaning and also reduces rotational ambiguity. Nevertheless, these problems are known to result in permutation, intensity and rotational ambiguity which may result in non-unique solutions. Another way to estimate kinetic parameters from spectroscopic data is based on a hard-modeling approach that also includes information from the kinetic model. For this method, this

model based on differential equations is integrated for fixed parameters. Then the concentration profiles are used to identify the absorbance profiles via least squares regression. The Gauss-Newton-Levenberg-Marquardt method is then used to update the values of the kinetic parameters based on the sensitivities with respect to the kinetic parameters and with respect to the improvement of the fit.

Apart from using soft- or hard-modeling techniques, hybrid techniques can be used, e.g. where MCR-ALS (MCR-Alternating Least Squares) is combined with a kinetic model description[6]. This methodology is widely used in academia but due to its challenges related to large datasets it is not widely used for industrial applications[7].

Moreover, Alsmeyer et al.[8] introduced an indirect spectral hard modeling approach which uses parameters for the descriptive models describing the individual spectra. This approach requires that all the individual absorbance profiles are known in advance and thus, cannot describe the behavior of intermediates or non-absorbing species. An alternative to the latter approach is proposed by Neymeyr et al.[9]. This approach is called pure component decomposition (PCD). It first uses singular value decomposition (SVD) to derive a low-rank approximation of the spectroscopic data matrix and then introduces a rotation matrix. As an objective the difference between the data matrix $\mathbf{D}$ and its model prediction is minimized. Non-negativity is ensured by adding penalty terms in the objective function. Sawall et al.[10] extended the PCD approach by combining it with a physical model that describes the reaction network and by adding a regularization term related to the concentration profiles to the objective function. This method obtains kinetic parameter estimates and the rotation matrix simultaneously but it is a multi-objective problem where many weighting factors have to be determined.

All of these sequential approaches can perform poorly for instable or ill-conditioned systems due to linearly dependent columns in the concentration matrix. Moreover, these methods do not include model noise and do not guarantee convergence to a solution.

To address these shortcomings, Chen et al.[3] developed a new methodology that proposes

a better alternative. This approach formulates a simultaneous optimization problem that results in kinetic parameter estimates and pure component spectra estimates simultaneously. The approach is introduced in more detail in Section 2. It has been tested on several case studies with simulated and real data from chemical and pharmaceutical industrial processes. To make the methodology accessible for industrial and academic researchers it has been implemented in a toolkit called *KIPET* (Kinetic Parameter Estimation Toolkit, Short et al.[1], Schenk et al.[2]) which we will employ for the investigations of the specific drug manufacturing process case study in this paper.

Drug manufacturing processes can be complex with multi-components and multiple sequence of operations, all of which can pose challenges for modeling. They frequently involve species of varying phases (liquid, solid, gas) reacting at different temperatures and often combined with several unit operations (reactions, extractions, additions, distillations etc.). For processes involving liquids and solids, certain assumptions need to be made to include the behavior of the solid in the reaction model. Some approaches have been made to relate the change in surface area to the reactant conversion, e.g. Grénman et al.[11], Salmi et al.[12]. However, this requires the particle behavior to be modeled, and most of these approaches cannot be implemented easily with complex reaction networks. An overview of some models and examples for their application can be found in Grénman et al.[11]. Gao[13] proposed a dissolution rate model describing the dissolution process by including the diffusion process at the solid-liquid interface. This approach includes the solid-liquid interface kinetics and mass transport kinetics in a flexible manner which makes it more applicable for data fitting purposes. Other modeling approaches have been made to model the dissolution process in a stochastic fashion. Lánský and Weiss[14] studied a heterogeneous tube model where the authors modeled the fractional dissolution via a deterministic model on the macroscopic level using stochastic modeling techniques at the microscopic level. Later Lánský and Weiss[15] applied this approach to general drug dissolution processes with homogeneous and heterogeneous particles using probabilistic concepts as well as random effects in dissolution.

Lánský and Weiss[16] studied a variety of empirical and semi-empirical models including homogeneous dissolution models especially with respect to the role of heterogeneity in drug dissolution models. Kalampokis et al.[17] modeled the fraction of dose absorbed and monitor it as a function of time using Monte Carlo methods. All these approaches bring along a set of additional probabilistic parameters that have to be estimated as well or additional Monte Carlo Sampling, which may be computationally expensive. Finally, Wang et al.[18] proposed a batch process model including a shrinking particle model for parameter estimation from concentration data using Bayesian estimation in order to investigate model discrimination. However, all of the approaches described above have not been used for parameter estimation from spectroscopic data so far. Here, we present a simpler, equilibrium-based macroscopic modeling approach that can also be used for parameter estimation from spectroscopic data. In the next section we present the kinetic model for the pharma case study, and describe the underlying kinetic parameter estimation problem and solution procedure. In section 3, the introduced methods are applied to the case study and results are discussed. The results for the first stage of the multi-stage process are presented. Subsequently, those results are used to initialize the second stage and the corresponding results for the second stage are described in detail. Both first and second stage results are then compared and discussed. Finally, section 4 concludes the study and presents directions for future work.

2 Model Formulation and Methodology

2.1 Process Background

In this pharmaceutical process case study, the synthesis of an unsymmetrical urea compound using carbonyldiimidazole (CDI) is investigated. The chemical reaction network is illustrated in fig. 1. The reaction is divided into two steps, [1] reacting CDI, the primary amine (S1), and an activating species (for simplicity referred to as compound X) in tetrahydrofuran (THF) to generate the acyl intermediate at low temperature, followed by [2] reaction with the second

primary amine (S2) to form the desired product (P) at higher temperature. During the first step, a symmetric urea dimer ($\tilde{D}$) can form through an isocyanate species, which is of concern because of a difficult purge in the crystallization. To aid process design and optimization, a kinetic model of this reaction sequence is highly desired.

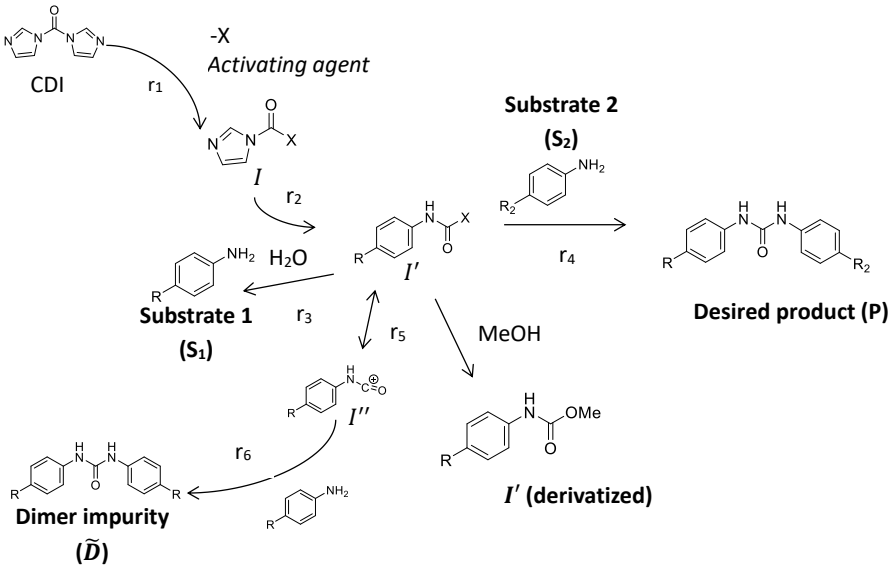

Figure 1: Reaction sequence used in the synthesis of the unsymmetrical urea product (P).

2.2 Model Derivation

The considered drug manufacturing process is described by the following chemical reaction network, where reaction 5 is an equilibrium reaction.

1. $\mathrm{CDI} + \mathrm{X} \xrightarrow{k_1} \mathrm{I}$
2. $\mathrm{I} + \mathrm{S_1} \xrightarrow{k_2} \mathrm{I'}$
3. $\mathrm{I'} + \mathrm{H_2O} \xrightarrow{k_3} \mathrm{S_1}$
4. $\mathrm{I'} + \mathrm{S_2} \xrightarrow{k_4} \mathrm{P}$
5. $\mathrm{I'} \underset{k_5^-}{\overset{k_5^+}{\rightleftharpoons}} \mathrm{I''} + \mathrm{X^-}$

6. $\mathrm{I''} \quad + \quad \mathrm{S_1} \xrightarrow{k_6} \tilde{\mathrm{D}}$

The full model based on differential algebraic equations (DAE) derived from the reaction network is given by:

$$
\begin{aligned}
&\frac{d}{dt}c_X = -r_1 + (r_5^+ - r_5^-)\left(\frac{1}{1+Ka\beta}\right), && c_X(0) = c_{X0}\\
&\frac{d}{dt}c_I = r_1 - r_2, && c_I(0) = c_{I0}\\
&\frac{d}{dt}c_{I'} = r_2 - r_3 - r_5^+ + r_5^- - \nu_3(t)r_4, && c_{I'}(0) = c_{I'0}\\
&\frac{d}{dt}c_{H_2O} = -r_3, && c_{H_2O}(0) = c_{H_2O0}\\
&\frac{d}{dt}c_{I''} = -r_6 + r_5^+ - r_5^-, && c_I''(0) = c_{I''0}\\
&\frac{d}{dt}c_{\tilde{D}} = r_6, && c_{\tilde{D}(0)} = c_{\tilde{D}0}\\
&\frac{d}{dt}\bar{c}_{CDI} = -\nu_1(t)r_1, && \bar{c}_{CDI}(0) = \bar{c}_{CDI_0}\\
&\frac{d}{dt}\bar{c}_{S_1} = \nu_2(t)(-r_2 + r_3 - r_6), && \bar{c}_{S_1}(0) = \bar{c}_{S_{10}}\\
&\frac{d}{dt}c_{S_2} = -\nu_3(t)r_4, && c_{S_2}(0) = c_{S_20}\\
&\frac{d}{dt}c_P = \nu_3(t)r_4, && c_P(0) = c_{P0}
\end{aligned}
$$

with the rate laws

$$
\begin{aligned}
r_1 &= k_1 c_{CDI} c_X\\
r_2 &= k_2 c_I c_{S_1}\\
r_3 &= k_3 c_{I'} c_{H_2O}\\
r_4 &= k_4 c_{I'} c_{S_2}\\
r_5^+ &= k_5^+ c_{I'}\\
r_5^- &= (\ k_5^-\ Ka\beta\ ) c_{I''} c_X\\
r_6 &= k_6 c_{I''} c_{S_1}.
\end{aligned}
$$

where k_i, $i = 1, \ldots, 6$ are rate constants and $Ka\beta$ is an aggregated equilibrium constant, which comes from the weak acid base equilibria, i.e.

$$Ka = \frac{c_{H^+} c_{X^-}}{c_X} \tag{2}$$

for a weak acid c_{H^+} with conjugate base c_{X^-}, such that

$$Ka\beta := \frac{Ka}{\beta} = \frac{c_{X^-}}{c_X} \tag{3}$$

and with this

$$\beta = \frac{c_X Ka}{c_{X^-}}. \tag{4}$$

The main challenge with this particular process is that it is a heterogeneous process with liquids and solids involved. Two solids are added at the start of the manufacturing process but there is no information available on which species is depleted first and at which point in time they are depleted. The dissolution or switching behavior from solid to liquid is modeled by the introduction of two switching functions $\nu_1(t)$ and $\nu_2(t)$ with $\nu_1 : \mathbb{R}^+ \to \{0, 1\}$ and $\nu_2 : \mathbb{R}^+ \to \{0, 1\}$, such that

$$\begin{aligned}
&\text{Switching for } CDI\text{:} \begin{cases} c_{CDI}(t) = c_{CDI}^{sat}(1 - \nu_1(t)) + \nu_1(t)\bar{c}_{CDI}(t) \\ \text{if } m_{CDI} = m_{CDI}^0 - \eta_1 \geq 0 : \ \nu_1(t) = 0, \\ \text{else } \ \nu_1(t) = 1, \end{cases} \\
&\text{Switching for } S_1\text{:} \begin{cases} c_{S_1}(t) = c_{S_1}^{sat}(1 - \nu_2(t)) + \nu_2(t)\bar{c}_{S_1}(t) \\ \text{if } m_{S_1} = m_{S_1}^0 - \eta_2 + \eta_3 - \eta_6 \geq 0 : \ \nu_2(t) = 0, \\ \text{else } \ \nu_2(t) = 1. \end{cases}
\end{aligned} \tag{5}$$

The reaction extent η_1 describes how much CDI has been consumed in reaction 1. In the same fashion the reaction extents η_2, η_6 and η_3 describe how much S_1 has been consumed in

reactions 2 and 6 and generated in reaction 3. The constants c_{CDI}^{sat} and $c_{S_1}^{sat}$ are the saturation concentrations for CDI and S_1 which are assumed to be at the solubility value. The differential variables $\bar{c}_{CDI}$ and $\bar{c}_{S_1}$ describe the behavior of CDI and S_1 in the liquid state when no solid CDI or S_1 remains. Note that IR spectroscopic data can only be obtained for measurements in the liquid phase. So we have to model the solid/liquid equilibrium in an appropriate manner.

This system is a multi-stage process where the first stage considers all the named reactions apart from reaction 4. Reaction 4 requires the addition of S_2, which is added at the beginning of the second stage, and leads to the product formation. This reaction is switched on or off with the function $\nu_3(t)$ with $\nu_3 : \mathbb{R}^+ \to \{0, 1\}$.

The switching functions $\nu_1(t)$, $\nu_2(t)$ and $\nu_3(t)$ introduce binary variables at every time point and turn this problem into a mixed-integer optimization problem. To avoid the large number of binary decisions, we consider a continuous representation of the switching functions instead. We still separate the two reaction stages so that $\nu_3(t)$ remains as a (fixed) switching decision, but we approximate $\nu_1(t)$ and $\nu_2(t)$ by the following sigmoidal functions, which we call smoothing functions. These sigmoidal functions come from the smoothed derivative of a step function.

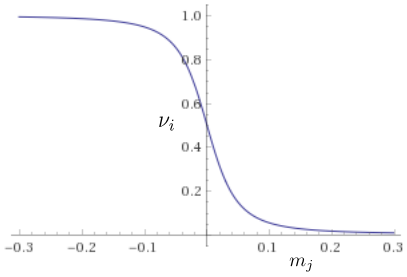

Figure 2: Smoothing ν_i vs. m_j with $i \in \{1, 2\}$ and $j \in \{CDI, S_1\}$, for $\epsilon = 5 \times 10^{-6}$, $\delta = 10^{-3}$ and $M = 10^{-8}$ (generated by Wolfram|Alpha[19])

$$\begin{aligned} \nu_1(t) &:= \frac{1}{2} + \frac{(\delta - m_{CDI})}{2((\delta - m_{CDI})^2 + \frac{\epsilon^2}{M})^{1/2}} \\ \nu_2(t) &:= \frac{1}{2} + \frac{(\delta - m_{S_1})}{2((\delta - m_{S_1})^2 + \frac{\epsilon^2}{M})^{1/2}} \end{aligned} \tag{6}$$

where we choose constants ϵ and δ with $\epsilon = 5 \times 10^{-6}$, $\delta = 1 \times 10^{-3}$ and M is the sigmoidal tuning parameter with $M \in [10^{-8}, 10^{-7}]$. For $M = 10^{-8}$, this function is illustrated for $m_j \in [-0.3, 0.3]$ and $j \in \{CDI, S_1\}$ in fig. 2. These smoothing functions depend on m_j with $j \in \{CDI, S_1\}$, to determine how much solid remains. We calculate this through the mass balances, and we add the following algebraic equations to our model

$$
\begin{aligned}
X &= X_0 - \eta_1 + \eta_5 \left(\frac{1}{1 + Ka\beta} \right) \\
I &= I_0 + \eta_1 - \eta_2 \\
H &= H_0 - \eta_3 \\
I^{''} &= I_0'' + \eta_5 - \eta_6 \\
\tilde{D} &= \tilde{D}_0 + \eta_6,
\end{aligned}
$$

where η_i with $i \in \{1, 2, 3, 5, 6\}$ denote our additional algebraic variables. Since we deal with the two stages separately, we specify the input $\nu_3(t)$, such that

$$
\nu_3(t) = \begin{cases} 0, & t \in \mathcal{T}_1 = [0.51667, 4.55] \\ 1, & t \in \mathcal{T}_2 = [4.55, 9], \end{cases}
$$

where $\mathcal{T}_1$ and $\mathcal{T}_2$ represent the first and second stages that are determined by the experimental times, respectively, of the process (in hours). That means ν_3 turns off the right hand side of the differential equations for c_{S_2} and c_P.

2.3 Kinetic Parameter Estimation Problem

All the examples in this study were performed using *KIPET*[1,2], where several contributions to the development of *KIPET* were also made. *KIPET* is an open-source toolkit written in Python with the algebraic modeling package *Pyomo*[20]. The software package makes use of the model formulations and concepts developed in Chen et al.[3]. The effectiveness of

the mathematical formulations and concepts has been shown for several problems, including many components and reactions and large real and simulated datasets. In this section we review the most important features of the underlying optimization problem but for the algorithms and methods implemented in the toolkit we refer to the original work by Chen et al. [3] or the *KIPET* related work by Short et al. [1] and Schenk et al. [2]. For the implementation within *KIPET* the spectral data has to follow Beer-Lambert's law eq. (1), and the reaction system here is a well-mixed fed-batch reaction system. Beer-Lambert's law describes the relationship between the pure components and their absorbance. Rewriting Beer-Lambert's law introduced in eq. (1) in scalar form results in

$$d_{i,l} = \sum_{k=1}^{nc} c_k(t_i) s_k(\lambda_l) + \zeta_{i,l}, \qquad i = 1, .., ntp, \quad l = 1, .., nwp \tag{7}$$

where $d_{i,l}$ is the spectroscopic data obtained at a sampling time t_i and wavelength λ_j, $c_k(t_i)$ is the concentration of species k at sampling time t_i and $s_k(\lambda_l)$ is the absorbance of species k at wavelength λ_j. ntp and nwp denote the total number of time points and wavelengths, and $\zeta_{i,l}$ is the measurement error.

To address the limitations of multivariate curve resolution techniques, Chen et al. [3] proposed a simultaneous approach that solves the full parameter estimation problem within an optimization framework, which includes eq. (7) along with a kinetic model of the form

$$\begin{aligned} &\frac{d\boldsymbol{z}(t)}{dt} = \boldsymbol{f}(\boldsymbol{z}(t), \boldsymbol{y}(t), \boldsymbol{\theta}), \\ &\boldsymbol{g}(\boldsymbol{z}(t), \boldsymbol{y}(t)) = 0, \\ &\boldsymbol{z}(t_0) = \boldsymbol{z_0}. \end{aligned} \tag{8}$$

Here $\boldsymbol{z}(t) \in \mathbb{R}_+^{nc}$ is the vector of concentrations coming from the kinetic model with nc as the total number of components and $\boldsymbol{y}(t) \in \mathbb{R}^{na}$ is the vector of algebraic variables with na as the total number of algebraic variables. Moreover, $\boldsymbol{z_0} \in \mathbb{R}_+^{nc}$ is the vector of initial concentrations, and $\boldsymbol{\theta} \in \mathbb{R}^{n\theta}$ are $n\theta$ estimated kinetic parameters. The kinetic model

functions $\boldsymbol{f} : \mathbb{R}^{nc+ny+n\theta} \rightarrow \mathbb{R}^{nf}$ with nf as the number of right hand sides of the differential equations and $\boldsymbol{g} : \mathbb{R}^{nc+ny+n\theta} \rightarrow \mathbb{R}^{ng}$ with ng as the number of algebraic equations, are assumed to be twice continuously differentiable with respect to the optimization variables $(\boldsymbol{z}, \boldsymbol{y}, \boldsymbol{\theta})^T$. As the predictions of eq. (8) are subject to model error, Chen et al.[3] proposed the following relationship between the predicted concentration $\boldsymbol{z}(t)$ and the concentration observed experimentally

$$c_k(t_i) = z_k(t_i) + \omega_k(t_i), \quad i = 1, \dots ntp, \quad k = 1, \dots, nc, \tag{9}$$

where ω_k denotes the model error. Combining eq. (7), eq. (8) and eq. (9) yields the following DAE-constrained optimization problem

$$\begin{aligned}
\min \quad & \frac{1}{\delta^2} \sum_{i=1}^{ntp} \sum_{l=1}^{nwp} \left(d_{i,l} - \sum_{k=1}^{nc} c_k(t_i) s_k(\lambda_l) \right)^2 + \sum_{i=1}^{ntp} \sum_{k=1}^{nc} \frac{1}{\sigma_k^2} (c_k(t_i) - z_k(t_i))^2 \\
\text{s.t.} \quad & \frac{d\boldsymbol{z}(t)}{dt} = \boldsymbol{f}(\boldsymbol{z}(t), \boldsymbol{y}(t), \boldsymbol{\theta}) \\
& \boldsymbol{g}(\boldsymbol{z}(t), \boldsymbol{y}(t)) = 0 \\
& \boldsymbol{z}(t_0) = \boldsymbol{z_0} \\
& \boldsymbol{c}(t_i) \geq 0 && i = 1, .., ntp \\
& \boldsymbol{s}(\lambda_l) \geq 0 \qquad (optional) && l = 1, .., nwp.
\end{aligned} \tag{P}$$

Throughout this study this optimization problem is referred to as Problem (P) and it is highlighted as the main formulation solved for this case study and solved within *KIPET*. In more detail, Problem (P) incorporates the kinetic model eq. (8) in the constraints while optimizing measurement and model error in its objective function. The first term in the objective minimizes deviations from Beer-Lambert's law and the second term minimizes deviations from the kinetic model. That means that $\zeta_{i,l}$ and ω_k are minimized implicitly. The variables in this optimization problem are the kinetic parameters $\boldsymbol{\theta}$, the species' unnoised and noised concentrations, $\boldsymbol{z}(t)$ and $\boldsymbol{c}(t_i)$ respectively, the species' absorbance profiles $\boldsymbol{s}(\lambda_l)$,

and the algebraic variables $\boldsymbol{y}(t)$. The degrees of freedom of Problem (P) are determined by the kinetic parameters. The spectral data, the $d_{i,l}$ values, are our input data considered in Problem (P). Furthermore, the variances δ^2 and σ_k^2 have to be known in order to solve Problem (P). The spectra data are available as measurements from experiments, although in general the variances are not. Therefore, to solve Problem (P), we first have to estimate these variances. This is realized using an iterative optimization-based heuristic according to Chen et al.[3].

2.4 Variance Estimation Procedure

Before solving Problem (P) the variances δ^2 and σ^2 need to be estimated. Within *KIPET* two methodologies for their estimation exist. The first was introduced in Chen et al.[3] and its functionality within *KIPET* demonstrated in Schenk et al.[2]. Recently Short et al.[21] proposed an alternative to the Chen et al.[3] approach, and this alternative is used to estimate the variances in this study. We summarize the procedure within the following lines and start with unknown covariances from two sample populations, i.e. instrumental measurement and modeling error. Then the joint probability distribution and likelihood function as well as log likelihood functions are derived. First we determine the upper bound v^2 for δ by setting $\sigma_k = 0$. Then we guess σ_k, solve problem (P) to determine $\delta^2 = \sum_i \epsilon_i^T \epsilon_i / (nwp\ ntp)$, and continue to update σ_k until the joint variance relation $f(\sigma_{k,p}) = v^2 - \delta^2 - (\sum_{l=1}^{nwp} \sum_{k=1}^{nc} \sigma_{k,p}^2 s_{kl}/nwp)$ is satisfied.

This variance estimation procedure has been shown to perform well, i.e. it is better behaved than the previous appoach and provides us with better data fits and kinetic parameter estimates. More details on this procedure including its performance for other case studies in comparison to the Chen et al.[3] approach can be found in Short et al.[21].

2.5 Estimation of Parameters and Determination of Confidence Regions

Now that we have a methodology to estimate the variances, we can investigate the solution of the main Problem (P). *KIPET* provides different ways for initialization of this problem and makes use of robust discretization methods based on orthogonal collocation on finite elements. It uses *IPOPT*[22] as the nonlinear programming (NLP) solver to compute a solution to this large-scale NLP. Within the KIPET framework the confidence regions of the estimated parameters can be determined using *sIPOPT*[23] or *k_aug*[24]. More details on the methodology behind *KIPET* and the general capabilities of the package can be found in Short et al.[1] and Schenk et al.[2].

2.6 Tuned Parameter Estimation Initialization Procedure

Due to the smoothing structure in eq. (6), the parameters have to be estimated step by step while refining the sigmoidal tuning parameter, in order to get a good initialization for the simultaneous parameter estimation problem and reveal the dependencies of the estimated parameter values on the considered tuning parameter value.

This initialization procedure is illustrated in fig. 3.

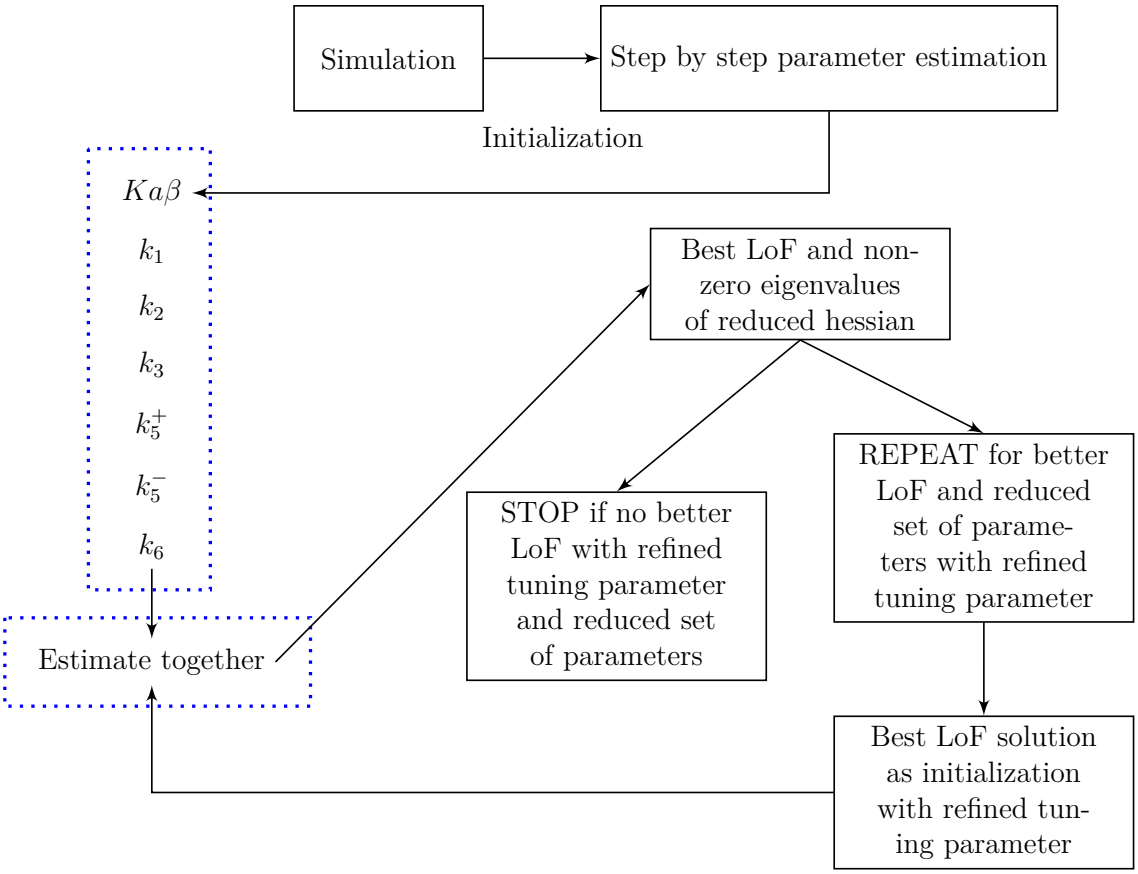

Figure 3: Tuned parameter estimation initialization procedure

We start with a simulation and an initial guess for the parameter vector θ, where $\theta = (Ka\beta, k_1, k_2, k_3, k_5^+, k_5^-, k_6)^T$. Then we initialize our first parameter estimation problem with this solution and estimate every parameter, one at a time. Our parameter estimation problem is of the form of Problem (P) with additional tight bounds for the parameters as

$$1.0 \geq Ka\beta \geq 0.9,$$

and

$$k_1, k_2, k_3, k_5^+, k_5^-, k_6 \geq 0.$$

The bounds for $Ka\beta$ result from the fact that we believe that c_X and c_{X^-} are almost the same. After the initial estimation from finding one parameter at a time, we check how many parameters we can estimate simultaneously, by initializing with the best Lack of Fit

(LoF) solution. The parameter $Ka\beta$ always goes to its upper bound, such that it has to be fixed with $Ka\beta = 1$. The simultaneous solution that results in the best LoF and low parameter variances for this tuning parameter is taken to initialize the new problem with increasing the value of the sigmoidal tuning parameter, M in eq. (6). We repeat this procedure until there is no improvement in LoF. If the estimated parameters are not at the bounds, we check the eigenvalues of the reduced Hessian. If they are close to zero, we fix the corresponding parameter at its current value. In this way we ensure tight parameter confidence intervals (CIs).

2.7 Lack of Fit as Estimation Quality Measure

The LoF is used as a relative metric of the quality of the parameters and model to fit the data to evaluate the performance[25], i.e.

$$LoF = \sqrt{\frac{\sum_{i=1}^{ntp}\sum_{l=1}^{nwp} e_{i,l}^2}{\sum_{i=1}^{ntp}\sum_{l=1}^{nwp} d_{i,l}^2}} \times 100\%, \tag{10}$$

where $e_{i,l}$ is the residual error corresponding to the corrected spectra element $d_{i,l}$.

2.8 Penalty Term Formulation

In some cases it can be useful to ensure that the estimated parameter values stay close to their nominal values. For this we introduce quadratic penalty terms to penalize values deviating from the previous estimates. This is added as a term in the objective of Problem (P), leading to Problem (P_{pen}).

$$
\begin{aligned}
\min \quad & \frac{1}{\delta^2}\sum_{i=1}^{ntp}\sum_{l=1}^{nwp}\left(d_{i,l}-\sum_{k=1}^{nc}c_k(t_i)s_k(\lambda_l)\right)^2+\sum_{i=1}^{ntp}\sum_{k=1}^{nc}\frac{1}{\sigma_k^2}(c_k(t_i)-z_k(t_i))^2+\sum_{m=1}^{n_{pen}}\omega_{pen}(\theta_m-\theta_m^\star)^2 \\
\text{s.t.} \quad & \frac{d\boldsymbol{z}(t)}{dt}=\boldsymbol{f}(\boldsymbol{z}(t),\boldsymbol{y}(t),\boldsymbol{\theta}) \\
& \boldsymbol{g}(\boldsymbol{z}(t),\boldsymbol{y}(t))=0 \\
& \boldsymbol{z}(t_0)=\boldsymbol{z_0} \\
& \boldsymbol{c}(t_i)\geq 0 \qquad\qquad i=1,..,ntp \\
& \boldsymbol{s}(\lambda_l)\geq 0 \qquad (optional) \qquad\qquad l=1,..,nwp,
\end{aligned}
\tag{P$_{pen}$}
$$

where ω_m is the penalty weighting parameter, $\theta_m^\star$ denotes the nominal value with $m = 1, \ldots, n_{pen}$, and n_{pen} is the number of parameters that should be ensured to stay close to the nominal value.

3 Results and Discussion

In this section the techniques introduced in the previous section are applied to the investigated drug manufacturing case study. We first investigate the first stage of the process and then initialize the second stage with those results and have a closer look at the second stage. All the numerical results presented in the following are generated on a 64bit Desktop PC with Ubuntu 16.04, with Intel(R) Core(TM) i7-7700 CPU at 3.60GHz with 16.3 GB of RAM.

3.1 First Stage Results

Typically, a data treatment step is advised to reduce the data to its most informative subset before using *KIPET*. For the first stage this turns out to be a wavenumber range of 1599 to 1851 which resulted from testing the informative quality of several ranges of wavenumbers. A principal component analysis (PCA) is performed on matrix $\mathbf{D}$, which

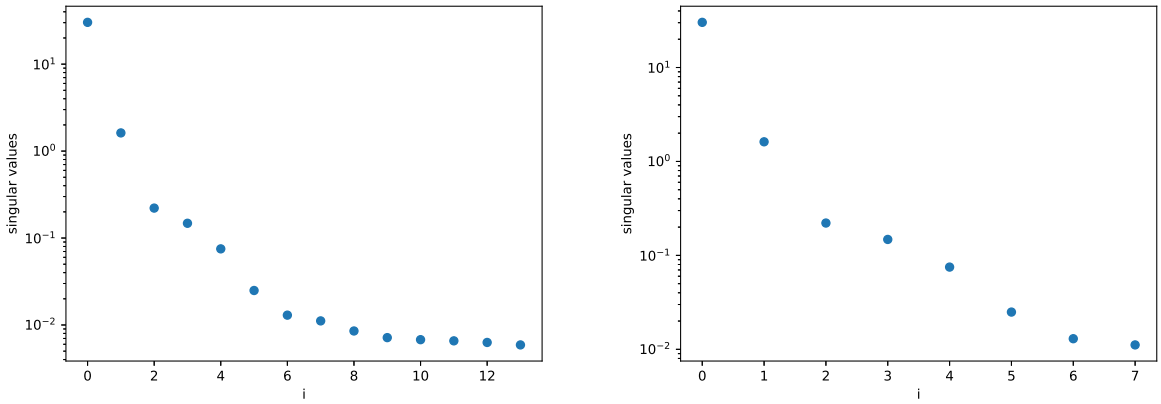

Figure 4: PCA results showing ranked singular values for 14 and 8 components

shows that only 6 out of 8 species should be assumed to be absorbing. The singular values in relationship to the components, resulting from a PCA assuming 14 principal components and 8 principal components, are illustrated in fig. 4. These figures show that the singular values for the 7th to 14th components hardly differ. The PCA for 8 principal components shows that there is a cut between the 6th and 7th component. That is why we assume 6 out of 8 species to be absorbing. As CDI and S_1 are mostly present as a solid in the first stage, we assume $\bar{c}_{CDI}$ and $\bar{c}_{S_1}$ to be non-absorbing.

The parameter estimation problem for this case study is the basic problem from Chen et al.[3] introduced in the last section. The results are generated with the following initial values obtained from the experimental conditions: $c_X = 0.344\,\mathrm{mol\,L^{-1}}, c_I = 0\,\mathrm{mol\,L^{-1}}, c_{I'} = 0\,\mathrm{mol\,L^{-1}}, c_{H_2O} = 0.0296\,\mathrm{mol\,L^{-1}}$(rounded to 4 digits), $c_{I''} = 0\,\mathrm{mol\,L^{-1}}, c_{\tilde{D}} = 0\,\mathrm{mol\,L^{-1}}, \bar{c}_{CDI} = 0.371\,\mathrm{mol\,L^{-1}}, \bar{c}_{S_1} = 0.082\,\mathrm{mol\,L^{-1}}$.

As described in the previous section, before we can solve the parameter estimation problem (P), we first estimate the variances. We assume that the system noise is the same for every component, select the tuning in fig. 2 with $M = 10^{-8}$ and then estimate the variances using the alternative variance estimation method described in Short et al.[21]. The resulting estimated variances are given in table 1. For the solution of the parameter estimation problems, we fix the variances to these values and solve Problem (P).

Table 1: Results for estimated variances using alternate variance estimation procedure

Error Source	Estimated Variance
Model (σ)	1.1125e-11
Instrument (σ)	1.0622e-05

Due to the smoothing structure, the parameters are estimated step by step while refining the tuning parameter as described in section 2.6 and illustrated in fig. 3. This procedure provides a good initialization for the simultaneous parameter estimation problem and reveals the dependencies of the estimated parameter values on the considered tuning parameter value.

3.1.1 Tuning Values of M

The results of the application of the tuned parameter estimation initialization procedure introduced in section 2.6 and highlighted in fig. 3 for this pharma process are now considered in more detail. We solve Problem (P) sequentially using different values for the sigmoidal tuning parameter M with $M \in \{10^{-8}, 2 \times 10^{-8}, 3 \times 10^{-8}, 5 \times 10^{-8}, 10^{-7}\}$ according to fig. 3. Increasing M after this leads to a worse LoF, such that the solution for $M = 10^{-7}$ is the final solution for the first stage.

Table 2: Parameter values and quality of estimates for different tuning parameters (e: estimated, f: fixed)

Tuning Parameter (M)/ Kinetic Parameters, Quality of Estimates	10^{-8}	2×10^{-8}	3×10^{-8}	5×10^{-8}	10^{-7}
$Ka\beta$	1.0 (f)	1.0 (f)	1.0 (f)	1.0 (f)	1.0 (f)
k_1	35.3705 (e)	40.4007 (e)	40.3513 (e)	40.4449 (e)	56.3128 (e)
k_2	7.6175 (e)	22.6608 (e)	24.3822 (e)	26.4930 (e)	66.1366 (e)
k_3	2.1273 (e)	0.9874 (e)	1.0228 (e)	1.0516 (e)	1.2973 (e)
k_5^-	907.5592 (f)	2622.0349 (e)	1411.7254 (e)	503.4825 (e)	1904.6545 (e)
k_5^+	0.0947 (e)	0.7473 (e)	0.5521 (e)	0.5264 (e)	0.4576 (e)
k_6	8.7720 (f)	8.7720 (f)	9.4422 (e)	6.3916 (e)	17.6065 (e)
LoF	0.6647%	0.5818%	0.5746%	0.5692%	0.5381%
max. CI	6.2%	10.8%	10.9%	8.1%	11.9%

Table 2 shows a comparison of the resulting kinetic parameter values, the LoFs and the

maximum percentile confidence intervals for the different tuning parameters. The parameter that varies the most with the refinement of the tuning parameter is k_5^-. However, with the last refinement 4 out of 6 parameters vary significantly.

The results for the final tuning parameter $M = 10^{-7}$ are presented below. The resulting

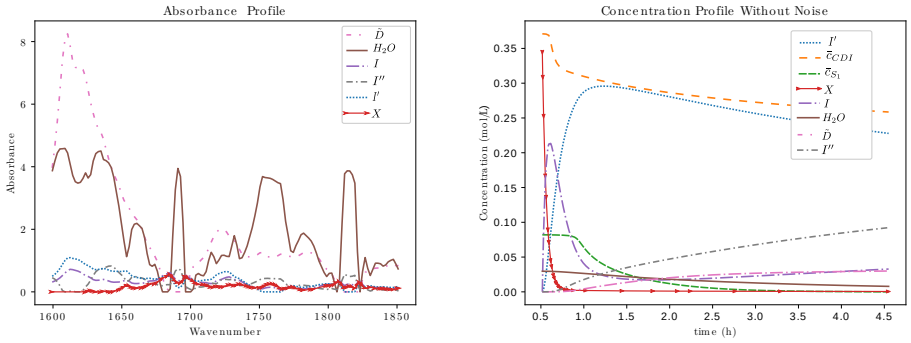

Figure 5: Absorbance and concentration profiles for $M = 10^{-7}$

absorbance and concentration profiles for this tuning parameter are illustrated in fig. 5. All species show absorbance activity and components $\tilde{D}$ and H show the most absorbance activity. The intermediates I, I' and I'' are formed, where I is almost immediately consumed again. The substrate X is consumed within the first hour. Figure 6 shows the depletion behavior and the smoothing function trajectories for $M = 10^{-7}$. The first graphic in fig. 6 shows the concentration profiles of the liquids and solids. Whenever the amounts of CDI and S_1 in solid phase are equal to zero the substrate just exists in liquid phase, such that m_{CDI} and m_{S_1} are computed for the whole first stage but the negative values can be neglected because they are dropped in the model formulation due to eq. (5). As for the previous tuning parameter values, the substrate CDI only stays at its solubility value for a very short amount of time. Then it is depleted and the differential formulation is used, as it is consumed in its liquid form. The substrate S_1 stays at its solubility value for the first hour until it is depleted and then consumed in its liquid form. With $M = 10^{-7}$ it is depleted earlier than for the previous tuning parameter values. The second graphic in fig. 6 shows the smoothing function trajectories, where both smoothing functions are much

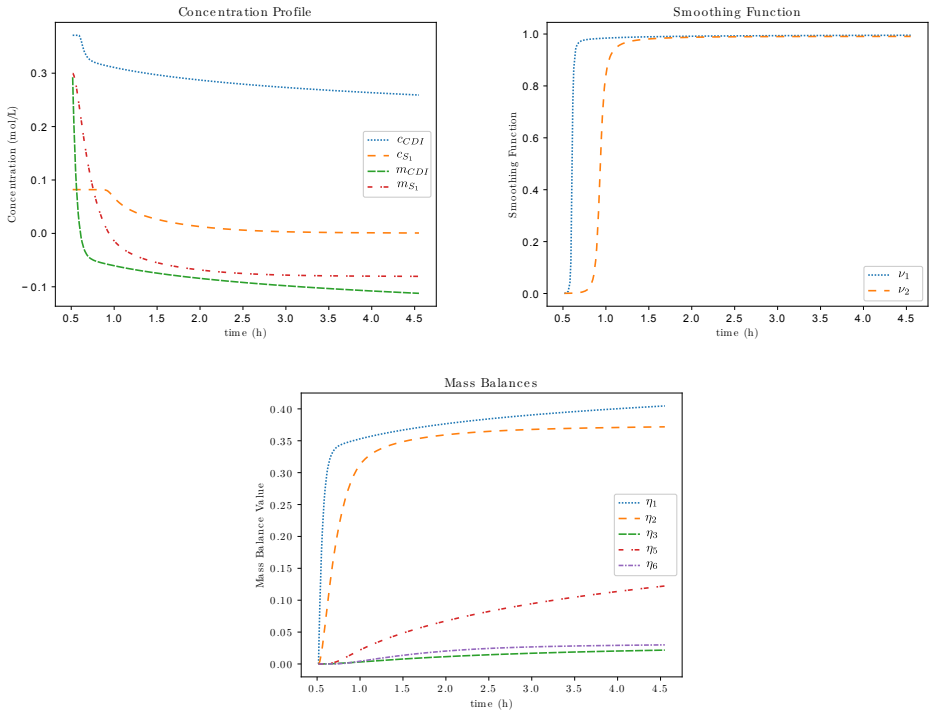

Figure 6: Concentration profiles for liquid and solid components and corresponding smoothing functions and mass balances for $M = 10^{-7}$

steeper than for the previous tuning parameter values and already give us good approximations for our switching functions. The third graphic in fig. 6 shows the mass balances that are used to estimate the amount of solids present.

When the set of simultaneously estimated parameters is reduced to 6 out of 7 parameters, the resulting LoF is 0.5381% and the estimated and fixed parameter values are

$$\begin{aligned}
Ka\beta &= 1.0 \\
k_1 &= 56.3128 \pm 1.3905 \quad 2.5\% \\
k_2 &= 66.1366 \pm 1.5867 \quad 2.4\% \\
k_3 &= 1.2973 \pm 0.0232 \quad 1.8\% \\
k_5^- &= 1904.6545 \pm 226.969 \quad 11.9\% \\
k_5^+ &= 0.4576 \pm 0.0317 \quad 6.9\% \\
k_6 &= 17.6065 \pm 1.1883 \quad 6.7\%
\end{aligned} \tag{11}$$

The quality of these estimates is very good. The same number of parameters as for the previous tuning parameter 5×10^{-8} can be estimated with similar confidence regions. Some parameters have larger confidence regions than in this previous case and others have tighter confidence regions. All estimated parameters vary by less than 12%. However, the smoothing functions give us much better approximations for this case than for the previous case. That is why we use these results to initialize the second stage.

3.2 Second Stage Results

In the first stage the reactor temperature was set to 0°C and raised to 50°C in the second stage. In a similar fashion as for the first stage, the IR dataset has to be reduced to a subset for the second stage. For the second stage this turned out to be a higher wavenumber range, i.e. 2442 to 2999.

As initial values for the differential components in the second stage, the final values from the first stage are taken, i.e.

$c_{H_2O} = 0.0078\,\mathrm{mol\,L^{-1}}, c_{\tilde{D}} = 0.0299\,\mathrm{mol\,L^{-1}}, c_I = 0.0328\,\mathrm{mol\,L^{-1}}, c_{I''} = 0.0923\,\mathrm{mol\,L^{-1}}, c_X = 0.0005\,\mathrm{mol\,L^{-1}}, c_{I'} = 0.2278\,\mathrm{mol\,L^{-1}}, \bar{c}_{CDI} = 0.2586\,\mathrm{mol\,L^{-1}}$ and $\bar{c}_{S_1} = 0\,\mathrm{mol\,L^{-1}}$ (rounded to 4 digits).

The initial conditions for the additional states S_2 and P are set to $c_{S_2}(4.55) = 68/250$ and $c_P(4.55) = 0.0$, where the second stage starts at 4.55 hours. It is assumed that $\bar{c}_{S_1}$, c_X, c_{H_2O}, c_P are non-absorbing with respect to the IR data because $\bar{c}_{S_1}(4.55)$, $c_X(4.55)$, $c_{H_2O}(4.55)$ are close to zero or equal to zero and for P we assume that it immediately crystallizes due to low solubility in the solvent mixture. This makes the estimation of the kinetic parameter k_4 a challenging task.

Before we solve the parameter estimation problem, we first estimate the variances, using the new variance estimation method in Short et al.[21] also used for the first stage. The resulting estimated variances are given in table 3.

For this case the switching functions (ν_1, ν_2, ν_3) are equal to one, and there is no need to

Table 3: Results for estimated variances using alternate variance estimation procedure

Error Source	Estimated Variance
Model (σ)	5.2553×10^{-13}
Instrument (δ)	1.5918×10^{-3}

tune M and apply eq. (6) as in the first stage. The parameter vector θ, now also includes k_4, such that $\theta = (Ka\beta, k_1, k_2, k_3, k_4, k_5^+, k_5^-, k_6)^T$ for the second stage. We assume that the temperature change mostly affects the product formation. Thus, we assume that the kinetic parameter estimates for all parameters apart from k_4 are very close to the estimated values in the first stage. This is why we introduce quadratic penalty terms to penalize values deviating from the previous estimates as introduced in section 2.8 and Problem (P_{pen}). However, some of the parameters have to be fixed as we obtain non-unique solutions and estimated parameters at the bounds, i.e. we fix k_2, k_3 and k_5^- to the previously estimated values given in eq. (11). We start with penalty weights for k_1, k_5^+, k_6 set to 10, 1 and 1. And then increase the weights to receive estimates closer to the previous ones. However, that worsens the LoF. This may indicate that the temperature dependence is not only reflected by the product formation parameter k_4 but also other kinetic parameters which are not included in the model, such that we keep the lower penalty weights. Modeling the temperature change would require estimating even more parameters using, e.g., an Arrhenius law formulation where also activation energies have to be estimated. However, for the set of parameters estimated here the data are not informative enough.

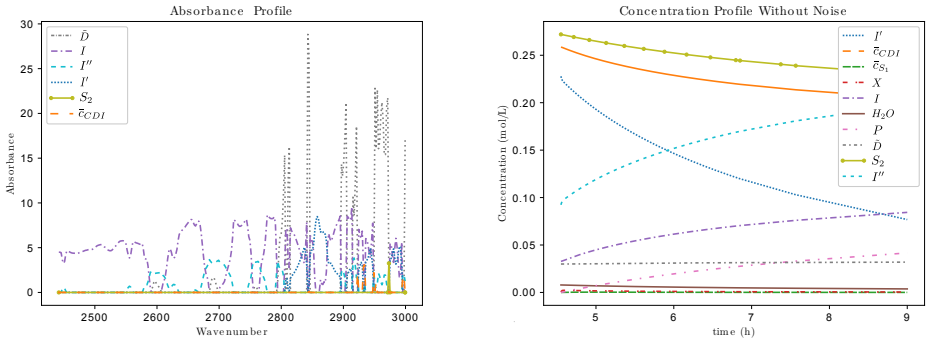

Figure 7: Absorbance and concentration profiles for IR data

Solving Problem (P_{pen}) with the first named penalty weights, i.e. 10, 1 and 1, we receive the absorbance and concentration profiles illustrated in fig. 7. We reduce the set of simultaneously estimated parameters to 4 out of 7 parameters and a penalty term is added for three of the four. The resulting LoF is 8.5300% for the IR data.
The estimated and fixed parameter values are

$$\begin{aligned}
Ka\beta &= 1.0 \\
k_1 &= 56.3564 \pm 0.0074 \quad 0.01\% \\
k_2 &= 66.1366 \\
k_3 &= 1.2974 \\
k_4 &= 0.2859 \pm 0.0064 \quad 2.2\% \\
k_5^- &= 1904.6545 \\
k_5^+ &= 2.1971 \pm 0.0202 \quad 0.9\% \\
k_6 &= 24.6215 \pm 0.0864 \quad 0.4\%.
\end{aligned} \tag{12}$$

The LoF for the second stage is worse than for the first stage but it is still below 9%. The estimates have good confidence regions; all are less than 3%.

4 Conclusions

This work addresses a number of challenges for kinetic model building in pharmaceutical processes. These deal with kinetic parameter estimation from spectroscopic data, but also include reactions with liquids and solids, where no measurement information is available about the depletion process. Moreover, we address large reactor temperature differences as well as multiple feeds during the process. This study develops a novel approach that addresses these challenges from the modeling side. All implementations were made in the open-source *KIPET*[1,2] package and the *KIPET* embedded concepts were further developed and used in the context of a particular drug manufacturing process case study. A tuned

parameter estimation initialization procedure was introduced that approximates the depletion process of the substrates that are present as solids at the beginning of the process. The variances, parameters, concentration and absorbance profiles were estimated for the first stage and then used to initialize the second stage. Subsequently, the variances, parameters, concentration and absorbance profiles for the second stage were estimated. In the first stage we found reliable estimates for 6 out of 7 kinetic parameters with very good confidence regions and small LoFs for both stages. Most parameter values for the second stage apart from k_4 use a quadratic penalty term in the objective to restrict deviations from the first stage estimated kinetic parameter values. The second stage resulted in an LoF that is less than 9% with confidence intervals less than 3%. This case study demonstrates that *KIPET* is a useful toolkit that deals with the rising challenges in kinetic model building. In particular, the outcomes of this work can be very useful for mathematical modeling of other multi-stage heterogeneous pharma processes with unknown solid depletion structure and multiple feeds.

5 Acknowledgements

The authors gratefully acknowledge the funding of Pfizer Inc.

References

(1) Short, M.; Schenk, C.; Thierry, D.; Rodriguez, J. S.; Biegler, L. T.; García-Muñoz, S. KIPET - An Open-Source Kinetic Parameter Estimation Toolkit. *Computer Aided Chemical Engineering, Proceedings of the 9th International Conference on Foundations of Computer-Aided Process Design* **2019**, *47*, 299–304, `https://www.sciencedirect.com/science/article/pii/B9780128185971500473`.

(2) Schenk, C.; Short, M.; Thierry, D.; Rodriguez, J. S.; Biegler, L. T.; García-Muñoz, S.;

Chen, W. Introducing KIPET : A Novel Open-Source Software Package for Kinetic Parameter Estimation from Experimental Datasets Including Spectra. *Computers and Chemical Engineering* **2020**, *134*, 106716, https://www.sciencedirect.com/science/article/abs/pii/S0098135419307811.

(3) Chen, W.; Biegler, L.; García-Muñoz, S. An approach for simultaneous estimation of reaction kinetics and curve resolution from process and spectral data. *Journal of Chemometrics* **2016**, *30*, 506–522, http://dx.doi.org/10.1002/cem.2808.

(4) Lawton, W. H.; Sylvestre, E. A. Self modeling curve resolution. *Technometrics* **1971**, *13*, 617–633.

(5) Lawton, W. H.; Sylvestre, E. A. Elimination of linear parameters in nonlinear regression. *Technometrics* **1971**,

(6) De Juan, A.; Maeder, M.; Martínez, M.; Tauler, R. Combining hard- and soft-modelling to solve kinetic problems. *Chemometrics and Intelligent Laboratory Systems* **2000**, *54*, 123–141.

(7) Chen, W.; Biegler, L.; García-Muñoz, S. Kinetic parameter estimation based on spectroscopic data with unknown absorbing species. *AIChE Journal* **2018**, *64*, 3595–3613, https://onlinelibrary.wiley.com/doi/abs/10.1002/aic.16334.

(8) Alsmeyer, F.; Koß, H.-J.; Marquardt, W. Indirect Spectral Hard Modeling for the Analysis of Reactive and Interacting Mixtures. *Applied Spectroscopy* **2004**, *58*, 975–985.

(9) Neymeyr, K.; Sawall, M.; Hess, D. Pure component spectral recovery and constrained matrix factorizations: Concepts and applications. *Journal of Chemometrics* **2010**, *24*, 67–74.

(10) Sawall, M.; Börner, A.; Kubis, C.; Selent, D.; Ludwig, R.; Neymeyr, K. Model-free multivariate curve resolution combined with model-based kinetics: Algorithm and applications. *Journal of Chemometrics* **2012**, *26*, 538–548.

(11) Grénman, H.; Salmi, T.; Murzin, D. Y. Studying Dissolution with a Model Integrating Solid-Liquid Interface Kinetics and Diffusion Kinetics. *Review Chemical Engineering* **2011**, 53–77.

(12) Salmi, T.; Grénman, H.; Wärnå, J.; Murzin, D. Y. New modeling approach to liquid-solid reaction kinetics: From ideal particles to real particles. *Chemical Engineering Research and Design* **2013**, *1*, 1876–1889.

(13) Gao, J. Y. Studying Dissolution with a Model Integrating Solid-Liquid Interface Kinetics and Diffusion Kinetics. *Analytical Chemistry* **2012**, *24*, 10671–10678.

(14) Lánský, P.; Weiss, M. Does the dose-solubility ratio affect the mean dissolution time of drugs? *Pharmaceutical Research* **1999**, *16*, 1470–1476, cited By 28.

(15) Lánský, P.; Weiss, M. Modeling Heterogeneity of Particles and Random Effects in Drug Dissolution. *Pharmaceutical Research* **2001**, *18*, 1061–1067, https://doi.org/10.1023/A:1010917118001.

(16) Lánský, P.; Weiss, M. Role of heterogeneity in deterministic models of drug dissolution and their statistical characteristics. *Biosystems* **2003**, *71*, 123 – 131, Selection of Papers from Topics in Biomathematics and Related Computational Problems at the Beginning of the Third Millennium, http://www.sciencedirect.com/science/article/pii/S0303264703001205.

(17) Kalampokis, A.; Argyrakis, P.; Macheras, P. A Heterogeneous Tube Model of Intestinal Drug Absorption Based on Probabilistic Concepts. *Pharmaceutical Research* **1999**, *16*, 1764–1769.

(18) Wang, Y.; Biegler, L. T.; Patel, M.; Wassick, J. Parameters estimation and model discrimination for solid-liquid reactions in batch processes. *Chemical Engineering Science* **2018**, *187*, 455 – 469, http://www.sciencedirect.com/science/article/pii/S0009250918303373.

(19) Wolfram|Alpha. Wolfram Alpha LLC, April 2019; https://www.wolframalpha.com/.

(20) Hart, W. E.; Laird, C. D.; Watson, J.-P.; Woodruff, D. L.; Hackebeil, G. A.; Nicholson, B. L.; Siirola, J. D. *Pyomo–Optimization Modeling in Python*, 2nd ed.; Springer Science & Business Media, 2017; Vol. 67.

(21) Short, M.; Biegler, L.; García-Muñoz, S.; Chen, W. Estimating Variances and Kinetic Parameters from Spectra Across Multiple Datasets Using KIPET. *Chemometrics and Intelligent Laboratory Systems* **2020**, 104012, https://doi.org/10.1016/j.chemolab.2020.104012.

(22) Wächter, A.; Biegler, L. T. On the Implementation of an Interior-Point Filter Line-Search Algorithm for Large-Scale Nonlinear Programming. *Mathematical Programming* **2006**, *106*, 25–57, http://dx.doi.org/10.1007/s10107-004-0559-y.

(23) Pirnay, H.; López-Negrete, R.; Biegler, L. T. Optimal Sensitivity based on IPOPT. *Mathematical Programming Computation* **2012**, *4*, 307–331, http://dx.doi.org/10.1007/s12532-012-0043-2.

(24) Thierry, D. M.; Biegler, L. T. Dynamic real-time optimization for CO_2 capture process. *AIChE J.* **2019**, *65*, 1–11, https://aiche.onlinelibrary.wiley.com/doi/abs/10.1002/aic.16511.

(25) De Juan, A.; Jaumot, J.; Tauler, R. Multivariate Curve Resolution (MCR). Solving the mixture analysis problem. *Analytical Methods* **2014**,