Incremental Parameter Estimation under Rank-Deficient Measurement Conditions

Kris Villez [1,2,*] **, Julien Billeter** [3] **and Dominique Bonvin** [3]

[1] Eawag, Swiss Federal Institute of Aquatic Science and Technology, 8600 Dübendorf, Switzerland
[2] ETH Zürich, Institute of Environmental Engineering, 8093 Zürich, Switzerland
[3] Laboratoire d'Automatique, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland; julien.billeter@alumni.epfl.ch (J.B.); dominique.bonvin@epfl.ch (D.B.)
[*] Correspondence: kris.villez@eawag.ch; Tel.: +41-58-765-5280

Abstract: The computation and modeling of extents has been proposed to handle the complexity of large-scale model identification tasks. Unfortunately, the existing extent-based framework only applies when certain conditions apply. Most typically, it is required that a unique value for each extent can be computed. This severely limits the applicability of this approach. In this work, we propose a novel procedure for parameter estimation inspired by the existing extent-based framework. A key difference with prior work is that the proposed procedure combines structural observability labeling, matrix factorization, and graph-based system partitioning to split the original model parameter estimation problem into parameter estimation problems with the least number of parameters. The value of the proposed method is demonstrated with an extensive simulation study and a study based on a historical data set collected to characterize the isomerization of α-pinene. Most importantly, the obtained results indicate that an important barrier to the application of extent-based frameworks for process modeling and monitoring tasks has been lifted.

Keywords: extents; graph theory; model identification; observability; optimal clustering; parameter estimation; state decoupling

1. Introduction

Despite advances in model identification theory, parameter estimation can still be very challenging in practice. Such challenges include the lack of identifiability, large computational cost, the need to formulate appropriate experimental designs, and the fact that many methods, such as those for uncertainty analysis, are still being investigated and therefore not standardized. In this work, we focus on a novel method to tackle the computational challenge associated with the identification of kinetic parameters in large dynamic models. Hence, the other challenges associated with parameter identifiability, experimental design, and uncertainty analysis, though relevant, are not investigated here.

To handle the computational challenge, it is typical to devise a protocol for model fitting and model validation. Such protocols can be divided into two broad classes. In the first class, the protocols are domain specific. For instance, several protocols for the identification of activated sludge models are discussed in [1]. Similarly, a protocol for environmental system models is proposed in [2]. These protocols incorporate significant expertise specific to the particular application domain, thereby leading to protocols that are fine-tuned for that domain. While they tend to be similar on a conceptual level, it is rather difficult to apply these protocols universally.

The second class includes protocols that are more general and thus—in principle—broadly applicable. The incremental model identification framework studied in [3–7] is a good example.

This method is grounded on the computation of *extents*, which are—loosely speaking—linear combinations of the original model states and capture the progress of individual dynamic phenomena, such as chemical reactions. A more precise definition will be given below.

The applicability of the extent-based framework is limited to cases where all extents can be computed based on measurements and invariant relationships. However, since this is rarely the case for biological process models, the extent-based framework has been extended in [8] to the case where each extent is either observable or non-sensed—see below for precise definitions. While this recent work provides a meaningful improvement, extent-based incremental model identification remains inapplicable for a wide range of biological scenarios found in practice. For example, this method cannot deal with the frequent case where an extent is sensed but unobservable.

The goal of this study is to present a novel method for incremental parameter identification that is more universally applicable. This method is based on the formulation of a generalized framework for extent computation and the use of a graph-based clustering algorithm. In what follows, we present the method and demonstrate its applicability to cases that could not be handled in an extent-based incremental model identification framework before. The expected impact of the newly developed tools is discussed at the end of this study.

2. Notation and Symbols

The matrix composed of the columns c of the matrix $\mathbf{M}$ is denoted $\mathbf{M}_{\bullet,c}$, while the matrix composed of the rows r of the matrix $\mathbf{M}$ is denoted $\mathbf{M}_{r,\bullet}$. All vectors are column vectors unless mentioned otherwise. Table 1 lists all symbols used in this study.

Table 1. List of symbols.

Symbol	Description	Dimensions
ϵ_h	Measurement error at time $t = t_h$	$M \times 1$
$\boldsymbol{\theta}\left(\boldsymbol{\theta}^{(j)}\right)$	Kinetic parameters (in subsystem j)	$T \times 1 \left(T^{(j)} \times 1\right)$
$\boldsymbol{\Lambda}_{\mathrm{o}}^{(j)}, \boldsymbol{\Lambda}_{\theta}^{(j)}$	Selection matrix	$\rho_{\mathrm{o}}^{(j)} \times \rho_{\mathrm{o}}, T^{(j)} \times T$
$\rho_{\mathrm{o}}\left(\rho_{\mathrm{o}}^{(j)}, \rho_{\mathrm{o,i}}^{(j)}\right)$	Number of observable extent directions (in system j, interpolated in system j)	1×1
ρ_{u}	Number of unobservable extent directions	1×1
ρ_{aug}	Number of extents and observable extent directions	1×1
$\boldsymbol{\Sigma}_{\epsilon}$	Measurement error variance-covariance matrix	$M \times M$
τ	Time (integrand)	1×1
$\boldsymbol{\Sigma}_{\overline{\chi}}$	Estimation error variance-covariance matrix	$A \times A$
χ_{aug}	Extents and observable extent directions	$\rho_{\mathrm{aug}} \times 1$
$\chi_{\mathrm{o}}\left(\chi_{\mathrm{o}}^{(j)}\right)$	Observable extent directions (in subsystem j)	$\rho_{\mathrm{o}} \times 1 \left(\rho_{\mathrm{o}}^{(j)} \times 1\right)$
$\overline{\chi}_{\mathrm{o}}\left(\overline{\chi}_{\mathrm{o}}^{(j)}\right)$	Observable extents and extent directions (in subsystem j)	$A \times 1 \left(A^{(j)} \times 1\right)$
$\tilde{\overline{\chi}}_{\mathrm{o},h}\left(\tilde{\overline{\chi}}_{\mathrm{o},h}^{(j)}\right)$	Computed observable extents and extent directions (in subsystem j)	$A \times 1 \left(A^{(j)} \times 1\right)$
$\tilde{\overline{\chi}}_{\mathrm{o,i}}\left(\tilde{\overline{\chi}}_{\mathrm{o,i}}^{(j)}\right)$	Interpolated observable extents and extent directions (in subsystem j)	$A \times 1 \left(A^{(j)} \times 1\right)$
χ_{u}	Unobservable extent directions	$\rho_{\mathrm{u}} \times 1$
$\overline{\chi}_{\mathrm{u}}$	Unobservable extents and extent directions	$(R - A) \times 1$
$A \left(A^{(j)}\right)$	Number of observable extents and extent directions (in subsystem j)	1×1
$\mathbf{a}$	Indices of ambiguous extents	$R_{\mathrm{a}} \times 1$
$\mathbf{B}$	Reduced row echelon form of $\mathbf{G}$	$M \times R$
$d_h\left(d_h^{(j)}\right)$	Model prediction residuals at t_h (in subsystem j)	$A \times 1 \left(A^{(j)} \times 1\right)$
$\mathbf{e}^{(j)}$	Indices of computed extents and extent directions simulated by subsystem j	$A^{(j)} \times 1$
$\mathcal{F}\left(\mathcal{F}^{(j)}\right)$	Information flow graph (for subsystem j)	$--$
$f\left(f^{(j)}\right)$	Rate expressions (in subsystem j)	$R \times 1 \left(R^{(j)} \times 1\right)$

Table 1. *Cont.*

Symbol	Description	Dimensions
G	Extent-based measurement matrix	$M \times R$
$\mathcal{G}_o, \overline{\mathcal{G}}_o$	Measurement matrix for the observable extent directions/observable extents and extent directions	$M \times \rho_o / A$
H	Number of measurements	1×1
h	Measurement sample index	1×1
i	Subsystem index	1×1
J	Number of subsystems	1×1
j	Indices of extents in subsystem j	$R^{(j)} \times 1$
j	Subsystem index	1×1
M	Species-based measurement matrix	$M \times S$
M	Number of measurement samples	1×1
N	Stoichiometric matrix	$R \times S$
$\overline{\mathcal{N}}_o, \overline{\mathcal{N}}_u$	Extent and extent direction-based stoichiometric matrices	$A/(R - A) \times S$
n	Indices of non-sensed extents	$R_n \times 1$
$\boldsymbol{n}\ (\boldsymbol{n}_0)$	Number of moles (at time $t = 0$)	$S \times 1$
$\mathbf{o}\left(\mathbf{o}^{(j)}\right)$	Indices of observable extents (in subsystem j)	$R_o \times 1\ \left(R_o^{(j)} \times 1\right)$
P	Projection matrix	$A \times M$
$Q_0, Q_0^*, Q_1^{(j)}$	Objective function	1×1
$R\ \left(R^{(j)}, R_a, R_n \atop R_o, R_o^{(j)}, R_{o,i}^{(j)}, R_s\right)$	Number of reactions (in subsystem j, ambiguous/non-sensed/ observable/observable in subsystem j/interpolated in subsystem j/sensed)	1×1
$\boldsymbol{r}$	Reaction rates	$R \times 1$
r	Reaction index	1×1
S	Number of species	1×1
s	Indices of sensed extents	$R_s \times 1$
$T\ (T^{(j)})$	Number of parameters (in subsystem j)	1×1
$t\ (t_h)$	Time (of measurement sample h)	1×1
$\mathbf{U}^{(j)}, \overline{\mathbf{U}}^{(j)}$	Mixing matrix	$\rho_o^{(j)}/A^{(j)} \times R^{(j)}$
$\mathbf{V}_o, \mathbf{V}_u$	Direction matrix for observable/unobservable direction directions	$R \times \rho_o, R \times \rho_u$
V	Reactor volume	1×1
$\mathbf{W}\left(\mathbf{W}^{(j)}\right)$	Weight matrix (for subsystem j)	$A \times A\ \left(A^{(j)} \times A^{(j)}\right)$
$\boldsymbol{x}\left(\boldsymbol{x}^{(j)}, x_r\right)$	Extents (in subsystem j, of reaction r)	$R \times 1\ \left(R^{(j)} \times 1, 1 \times 1\right)$
$\boldsymbol{y}\ (\boldsymbol{y}_0, \boldsymbol{y}(t_h))$	Noise-free measurements (at time $t = 0$, $t = h$)	$M \times 1$
$\tilde{\boldsymbol{y}}_h$	Measurements in sample h	$M \times 1$

3. Methods

3.1. System Representation and Extents

3.1.1. Dynamic Model in Terms of Numbers of Moles

We study batch process systems whose dynamic behavior is described by a set of differential equations of the following form:

$$\dot{\boldsymbol{n}}(t) = \mathbf{N}^{\mathrm{T}}\, V\, \boldsymbol{r}(t), \qquad\qquad \boldsymbol{n}(0) = \boldsymbol{n}_0. \qquad\qquad (1)$$

In the present work, the above equations represent a single-phase reaction system in a vessel with constant volume V. $\boldsymbol{n}(t)$ is the S-dimensional vector of numbers of moles; $\boldsymbol{n}_0$ specifies the initial conditions; $\boldsymbol{r}(t)$ is the R-dimensional vector of reaction rates; $\mathbf{N}$ is the $(R \times S)$-dimensional stoichiometric matrix.

We assume that M noisy measurements, $\tilde{\boldsymbol{y}}_h$, are available at the H sampling times t_h ($h = 1, \ldots, H$), according to the following equations:

$$\tilde{\boldsymbol{y}}_h := \boldsymbol{y}(t_h) + \boldsymbol{\epsilon}_h = \frac{1}{V} \mathbf{M}\, \boldsymbol{n}(t_h) + \boldsymbol{\epsilon}_h, \qquad\qquad \boldsymbol{\epsilon}_h \sim \mathcal{N}\left(\mathbf{0}, \boldsymbol{\Sigma}_\epsilon\right), \tag{2}$$

where $\boldsymbol{y}(t_h)$ is the M-dimensional vector of noise-free measurements, $\boldsymbol{\epsilon}_h$ the M-dimensional vector of measurement errors, $\boldsymbol{\Sigma}_\epsilon$ the M-dimensional measurement error variance-covariance matrix, and $\mathbf{M}$ is the $(M \times S)$-dimensional *species-based measurement matrix*.

We further assume that kinetic laws for the reaction rates are available. These laws are functions of the component masses, $\boldsymbol{n}(t)$, and a T-dimensional vector of kinetic parameters, $\boldsymbol{\theta}$:

$$\boldsymbol{r}(t) := \boldsymbol{f}\left(\boldsymbol{n}(t), \boldsymbol{\theta}\right). \tag{3}$$

In the present study, we assume that the rate laws are known, except for the values of the parameters $\boldsymbol{\theta}$ that need to be estimated.

3.1.2. Dynamic Model in Terms of Extents

We now adopt the definition of the *extents of reaction*, hereafter simply called *extents*, as the number of times each reaction has occurred since $t = 0$. These extents are measured in moles and defined mathematically as:

$$\boldsymbol{x}(t) := V \int_0^t \boldsymbol{r}\left(\tau\right) d\tau. \tag{4}$$

It follows that

$$\boldsymbol{n}(t) = \boldsymbol{n}_0 + \mathbf{N}^\mathsf{T}\, \boldsymbol{x}(t). \tag{5}$$

Accordingly, (1)–(3) can be represented equivalently in the following form:

$$\dot{\boldsymbol{x}}(t) = V\, \boldsymbol{f}\left(\boldsymbol{n}(t), \boldsymbol{\theta}\right), \qquad\qquad \boldsymbol{x}(0) = \mathbf{0} \tag{6}$$

$$\tilde{\boldsymbol{y}}_h = \frac{1}{V} \mathbf{M}\, \left(\boldsymbol{n}_0 + \mathbf{N}^\mathsf{T}\, \boldsymbol{x}(t_h)\right) + \boldsymbol{\epsilon}_h$$

$$= \boldsymbol{y}_0 + \frac{1}{V}\, \mathbf{G}\, \boldsymbol{x}(t_h) + \boldsymbol{\epsilon}_h \tag{7}$$

$$\boldsymbol{n}(t) = \boldsymbol{n}_0 + \mathbf{N}^\mathsf{T}\, \boldsymbol{x}(t), \tag{8}$$

where

$$\boldsymbol{y}_0 := \frac{1}{V}\, \mathbf{M}\, \boldsymbol{n}_0 \tag{9}$$

$$\mathbf{G} := \mathbf{M}\, \mathbf{N}^\mathsf{T}, \tag{10}$$

with the $(M \times R)$-dimensional matrix $\mathbf{G}$ being labeled the *extent-based measurement matrix*.

Let A denote the rank of $\mathbf{G}$, that is, $A := \mathsf{rank}\left(\mathbf{G}\right) \le \min\left(M, R\right)$.

3.2. Labeling Extents

3.2.1. Definitions

To label the extents, the following definitions are proposed:

- The rth extent $x_r(t)$ is labeled *sensed* if the measurements $\tilde{\boldsymbol{y}}_h$ are affected by $x_r(t)$ through the measurement Equation (7), that is, if $\mathbf{G}_{\bullet,r}$ has at least one non-zero element.
- The rth extent is labeled *observable* if (i) it is sensed ($\mathbf{G}_{\bullet,r}$ not null); and (ii) the change in the measurements $\tilde{\boldsymbol{y}}_h$ caused by a change in $x_r(t)$ can be unambiguously attributed to the change in that extent, that is, $\mathbf{G}_{\bullet,r}$ is independent of all other column vectors in $\mathbf{G}$.

The above labeling is based on the structure of **G** and does not depend on the temporal resolution or quality of the recorded measurements ($\tilde{y}_h$). As such, this means only *structural* observability is considered. This terminology is similar to the one used in data reconciliation [9,10] as the labels are produced without a dynamic model.

3.2.2. Labeling Procedure

To label the extents, one first computes the $(M \times R)$-dimensional reduced row echelon form of **G** denoted **B**:

$$\mathbf{B} := \mathrm{rref}\,(\mathbf{G}).\qquad(11)$$

This $(M \times R)$-dimensional matrix is composed of A non-zero rows and $M - A$ zero rows. Using **B**, the extents are labeled with the following procedure:

(a) Label the R_n extents corresponding to zero columns in **B** as *non-sensed* and use the vector **n** to identify their positions in x. Label the R_s remaining extents as *sensed* and use the vector **s** to identify their positions in x.

(b) Find all rows in **B** with a *single* non-zero element and find the column positions of these non-zero elements. Label the R_o extents corresponding to these column positions as *observable* and use the vector **o** to identify their positions in x. These extents are observable because one can compute a unique value based on the available information (measurements, extent-based measurement matrix, and initial conditions).

(c) Label the R_a extents that are sensed but not observable as *ambiguous*. These extents are ambiguous because one cannot compute a unique value based on the available information. The vector **a** is used to identify their positions in x.

Based on this labeling, the vectors $x_s(t)$, $x_n(t)$, $x_o(t)$, and $x_a(t)$ are defined to represent (i) the R_s sensed extents, (ii) the $R_n = R - R_s$ non-sensed extents; (iii) the R_o observable extents; and (iv) the $R_a = R_s - R_o$ ambiguous extents. This is illustrated in Figure 1.
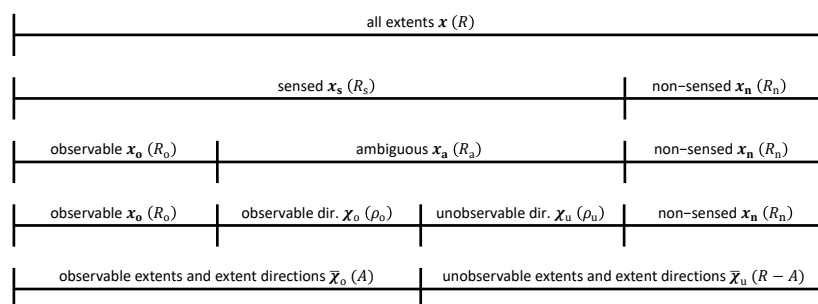

Figure 1. Labeling of the R extents as sensed, non-sensed, observable, and ambiguous. The ambiguous extents are spanned by ρ_o observable and ρ_u unobservable extent directions. This way, the extent space can be represented by $A = R_o + \rho_o$ observable extents and extent directions, and $R - A = R_n + \rho_u$ unobservable extents and extent directions.

With the above definitions, (5) can be reformulated as:

$$
\begin{aligned}
n(t) &= n_0 + \mathbf{N_{s,\bullet}}^{\mathrm{T}} x_s(t) + \mathbf{N_{n,\bullet}}^{\mathrm{T}} x_n(t)\\
&= n_0 + \mathbf{N_{o,\bullet}}^{\mathrm{T}} x_o(t) + \mathbf{N_{a,\bullet}}^{\mathrm{T}} x_a(t) + \mathbf{N_{n,\bullet}}^{\mathrm{T}} x_n(t).
\end{aligned}\qquad(12)
$$

Furthermore, it follows from $\mathbf{G_{\bullet,n}} = \mathbf{0}_{M \times R_n}$ that

$$\tilde{y}_h = y_0 + \frac{1}{V}\,\mathbf{G}_{\bullet,\mathbf{s}}\,x_\mathbf{s}(t_h) + \epsilon_h$$
$$= y_0 + \frac{1}{V}\,\mathbf{G}_{\bullet,\mathbf{o}}\,x_\mathbf{o}(t_h) + \frac{1}{V}\,\mathbf{G}_{\bullet,\mathbf{a}}\,x_\mathbf{a}(t_h) + \epsilon_h. \tag{13}$$

3.2.3. Practical Cases

Depending on the values of A, M, and R, one can distinguish the following cases:

I.   Full-rank extent-based measurement matrix ($A = R$). This occurs when there are at least as many measurements as reactions ($M \geq R$) and the matrix $\mathbf{G}$ is full column rank. As a result, $R_\mathbf{n} = 0$, $R_\mathbf{o} = R$, and $R_\mathbf{a} = 0$. This is the most frequently studied case, e.g., in [5–7]. This case enables computing unique values for the R extents by means of a linear transformation [5], thus allowing the estimation of kinetic parameters for each reaction rate model individually.

II. Rank-deficient measurement matrix ($A < R$). If $A < R$, for example because there are fewer measurements than reactions ($M < R$), it is no longer possible to compute all R extents from M measurements without additional information such as an established kinetic model. One can distinguish two situations within this case:

(a) No ambiguity ($R_\mathbf{a} = 0$). In this case, $R_\mathbf{o} = A$ and $R_\mathbf{o} + R_\mathbf{n} = R$. As shown in [8], it is possible in this case to implement efficient parameter estimation by identifying subsystems of the complete model that include a subset of the kinetic rate laws and their parameters.

(b) Ambiguity present ($0 < R_\mathbf{a} \leq A$). This situation results in $R_\mathbf{o} < A$. For this case, no generally applicable method for incremental parameter estimation is available until now.

The method proposed in this work is developed with the aim of handling all the above cases in a single framework for extent-based kinetic parameter estimation.

3.3. Observable and Unobservable Extent Directions

The ambiguous extents are now investigated in more detail to determine observable *directions* among them.

3.3.1. Factorization of $\mathbf{G}_{\bullet,\mathbf{a}}$

Let $\rho_\mathbf{o}$ be the rank of the ($M \times R_\mathbf{a}$)-dimensional measurement matrix $\mathbf{G}_{\bullet,\mathbf{a}}$. Then, $\mathbf{G}_{\bullet,\mathbf{a}}$ can be factorized into the ($M \times \rho_\mathbf{o}$)-dimensional measurement matrix $\mathcal{G}_\mathbf{o}$ and the ($R_\mathbf{a} \times \rho_\mathbf{o}$)-dimensional matrix $\mathbf{V}_\mathbf{o}$:

$$\mathbf{G}_{\bullet,\mathbf{a}} = \mathcal{G}_\mathbf{o}\,\mathbf{V}_\mathbf{o}^\mathsf{T}. \tag{14}$$

Using the reduced row echelon form $\mathbf{B}_{\bullet,\mathbf{a}}$, the matrices $\mathcal{G}_\mathbf{o}$ and $\mathbf{V}_\mathbf{o}$ are computed as follows:

- $\mathbf{V}_\mathbf{o}^\mathsf{T}$ is obtained by selecting the $\rho_\mathbf{o}$ non-zero rows in $\mathbf{B}_{\bullet,\mathbf{a}}$,
- $\mathcal{G}_\mathbf{o}$ is the matrix composed of the columns of $\mathbf{G}_{\bullet,\mathbf{a}}$ corresponding to the column positions of the first non-zero elements in the rows of $\mathbf{V}_\mathbf{o}^\mathsf{T}$.

3.3.2. Definition of Observable and Unobservable Extent Directions

The term $\mathbf{G}_{\bullet,\mathbf{a}}\,x_\mathbf{a}(t_h)$ in (13) can be expressed as:

$$\mathbf{G}_{\bullet,\mathbf{a}}\,x_\mathbf{a}(t) = \mathcal{G}_\mathbf{o}\,\mathbf{V}_\mathbf{o}^\mathsf{T}\,x_\mathbf{a}(t) = \mathcal{G}_\mathbf{o}\,\chi_\mathbf{o}(t), \tag{15}$$

with $\chi_{\mathrm{o}}(t)$ the ρ_{o}-dimensional vector of *observable extent directions* among the ambiguous extents:

$$\chi_{\mathrm{o}}(t) := \mathbf{V_o}^{\mathrm{T}} x_{\mathbf{a}}(t) = \mathbf{V_o}^{\mathrm{T}} \mathbf{I_{a,\bullet}} x(t), \tag{16}$$

where the $(R_{\mathbf{a}} \times R)$-dimensional matrix $\mathbf{I_{a,\bullet}}$ includes the rows $\mathbf{a}$ of the identity matrix $\mathbf{I}_R$ such that

$$x_{\mathbf{a}}(t) = \mathbf{I_{a,\bullet}} \, x(t). \tag{17}$$

Equation (15) indicates that while the extents $x_{\mathbf{a}}(t)$ cannot be observed individually, their combined effects on the measurements can be observed as the linear combinations $\chi_{\mathrm{o}}(t)$.

Remark 1. *The unobservable extent directions span the null space of* $\mathbf{G_{\bullet,a}}$, *which is also the null space of* $\mathbf{V_o}^T$. *Denoting this null space by the* $(R_{\mathbf{a}} \times \rho_{\mathbf{u}})$-*dimensional matrix* $\mathbf{V_u}$,

$$\mathbf{V_u} = \mathsf{null}\left(\mathbf{V_o}^T\right), \tag{18}$$

with $\rho_{\mathbf{u}} = (R_{\mathbf{a}} - \rho_{\mathrm{o}})$, *one can define the* $\rho_{\mathbf{u}}$-*dimensional vector of* unobservable extent directions $\chi_{\mathbf{u}}(t)$ *as:*

$$\chi_{\mathbf{u}}(t) := \mathbf{V_u}^T x_{\mathbf{a}}(t) = \mathbf{V_u}^T \mathbf{I_{a,\bullet}} \, x(t). \tag{19}$$

3.4. Observable Extents and Extent Directions

We further define the vector $\overline{\chi}_{\mathrm{o}}(t)$ consisting of the $A = R_{\mathrm{o}} + \rho_{\mathrm{o}}$ observable extents and extent directions as follows:

$$\overline{\chi}_{\mathrm{o}}(t) := \left[\begin{array}{c} x_{\mathbf{o}}(t) \\ \chi_{\mathbf{o}}(t) \end{array} \right] = \left[\begin{array}{c} x_{\mathbf{o}}(t) \\ \mathbf{V_o}^{\mathrm{T}} x_{\mathbf{a}}(t) \end{array} \right]. \tag{20}$$

With this definition, the measurement Equation (13) can be rewritten as:

$$\tilde{y}_h = y_0 + \frac{1}{V} \, \mathbf{G_{\bullet,o}} \, x_{\mathbf{o}}(t_h) + \frac{1}{V} \, \mathcal{G}_{\mathrm{o}} \, \chi_{\mathbf{o}}(t) + \epsilon_h$$

$$= y_0 + \frac{1}{V} \, \overline{\mathcal{G}}_{\mathrm{o}} \, \overline{\chi}_{\mathrm{o}}(t_h) + \epsilon_h, \tag{21}$$

with $\overline{\mathcal{G}}_{\mathrm{o}}$, an $(M \times A)$-dimensional matrix, constructed as

$$\overline{\mathcal{G}}_{\mathrm{o}} := \left[\begin{array}{cc} \mathbf{G_{\bullet,o}} & \mathcal{G}_{\mathrm{o}} \end{array} \right]. \tag{22}$$

Since the $(R_{\mathbf{a}} \times A)$-dimensional matrix $\overline{\mathcal{G}}_{\mathrm{o}}$ has full column rank, (21) can be used to compute the maximum-likelihood estimates $\tilde{\overline{\chi}}_{\mathrm{o},h}$ of the observable extents and extent directions directly from the measurements:

$$\tilde{\overline{\chi}}_{\mathrm{o},h} = \mathbf{P}\left(\tilde{y}_h - y_0\right), \tag{23}$$

with the $(A \times M)$-dimensional matrix $\mathbf{P}$ given by:

$$\mathbf{P} = V \left(\overline{\mathcal{G}}_{\mathrm{o}}^{\mathrm{T}} \, \Sigma_{\epsilon}^{-1} \, \overline{\mathcal{G}}_{\mathrm{o}} \right)^{-1} \overline{\mathcal{G}}_{\mathrm{o}}^{\mathrm{T}} \, \Sigma_{\epsilon}^{-1}. \tag{24}$$

The associated expected variance-covariance matrix of the estimation errors becomes:

$$\Sigma_{\overline{\chi}} = \mathbf{P} \, \Sigma_{\epsilon} \, \mathbf{P}^{\mathrm{T}} = \left(\overline{\mathcal{G}}_{\mathrm{o}}^{\mathrm{T}} \, \Sigma_{\epsilon}^{-1} \, \overline{\mathcal{G}}_{\mathrm{o}} \right)^{-1}. \tag{25}$$

3.5. Unobservable Extents and Extent Directions

We finally define the vector $\overline{\chi}_{\mathsf{u}}(t)$ consisting of the $R - A = R_{\mathsf{n}} + \rho_{\mathsf{u}}$ unobservable extents and non-sensed extent directions as follows:

$$\overline{\chi}_{\mathsf{u}}(t) := \begin{bmatrix} x_{\mathsf{n}}(t) \\ \chi_{\mathsf{u}}(t) \end{bmatrix} = \begin{bmatrix} x_{\mathsf{n}}(t) \\ \mathbf{V_u}^{\mathsf{T}} x_{\mathsf{a}}(t) \end{bmatrix} = \begin{bmatrix} \mathbf{I_{n,\bullet}} \\ \mathbf{V_u}^{\mathsf{T}} \mathbf{I_{a,\bullet}} \end{bmatrix} x(t). \tag{26}$$

With this definition, the expression for the number of moles (12) can be rewritten as:

$$
\begin{aligned}
n(t) &= n_0 + \mathbf{N_{o,\bullet}}^{\mathsf{T}} x_{\mathsf{o}}(t) + \mathbf{N_{a,\bullet}}^{\mathsf{T}} \left((\mathbf{V_o})^{+\mathsf{T}} \chi_{\mathsf{o}}(t) + (\mathbf{V_u})^{+\mathsf{T}} \chi_{\mathsf{u}}(t) \right) + \mathbf{N_{n,\bullet}}^{\mathsf{T}} x_{\mathsf{n}}(t) \\
&= n_0 + \overline{\mathcal{N}}_{\mathsf{o}}^{\mathsf{T}} \overline{\chi}_{\mathsf{o}}(t) + \overline{\mathcal{N}}_{\mathsf{u}}^{\mathsf{T}} \overline{\chi}_{\mathsf{u}}(t)
\end{aligned}
\tag{27}
$$

with:

$$\overline{\mathcal{N}}_{\mathsf{o}} := \begin{bmatrix} \mathbf{N_{o,\bullet}} \\ (\mathbf{V_o})^{+} \mathbf{N_{a,\bullet}} \end{bmatrix} \tag{28}$$

$$\overline{\mathcal{N}}_{\mathsf{u}} := \begin{bmatrix} \mathbf{N_{n,\bullet}} \\ (\mathbf{V_u})^{+} \mathbf{N_{a,\bullet}} \end{bmatrix} \tag{29}$$

3.6. System Partitioning

Incremental parameter estimation is based on the possibility of separating the parameter estimation problem into smaller problems. Ideally, if all extents of reaction could be computed from measurements, each reaction could be identified individually, that is, independently of the other reactions. This way, the parameter estimation problem reduces to the solution of R smaller problems. As discussed in the previous section, some extents may not be observable. For these situations, it would still be nice to be able to separate the parameter estimation problem into J smaller problems ($J \leq R$). The objective is therefore to partition the reaction system effectively—with as many small groups of reactions as possible—to simplify the parameter estimation task. The first step to achieve this consists of system partitioning.

An algorithm is developed to group the kinetic parameters into J parameter subsets ($j = 1, \ldots, J$), each represented as a $T^{(j)}$-dimensional vector $\boldsymbol{\theta}^{(j)}$ satisfying the following properties:

- The size $T^{(j)}$ of each parameter subset should be as small as possible.
- The estimates $\hat{\boldsymbol{\theta}}^{(j)}$ in the jth parameter subset can be computed without consideration of any other parameter subset $\boldsymbol{\theta}^{(i)}, i \neq j$.
- Each parameter in $\boldsymbol{\theta}$ appears in at most one of the parameter subsets $\boldsymbol{\theta}^{(j)}$.

This objective is achieved by means of model reformulation and graph-based system partitioning. The graph-based procedure can be interpreted as a symbolic manipulation of the process model. It does not require symbolic differentiation, however.

3.6.1. Step 1—Model Reformulation

An extended model is first defined to describe the dynamics of all extents and all observable directions among the ambiguous extents. To this end, the following procedure is applied:

(a) *Express $\dot{x}(t)$ as a function of $\overline{\chi}_{\mathsf{o}}(t)$ and $x(t)$*

The dynamic model (6) is modified by replacing $n(t)$ with the right-hand side of (27):

$$\dot{x}(t) = V f\left(\overline{\chi}_{\mathsf{o}}(t), \overline{\chi}_{\mathsf{u}}(t), \boldsymbol{\theta} \right), \qquad\qquad x(0) = \mathbf{0}. \tag{30}$$

The vector $\overline{\chi}_u(t)$ is now replaced with the right-hand side of (26). As a result, the above system becomes:

$$\dot{x}(t) = V f\left(\overline{\chi}_o(t), x(t), \theta\right), \qquad\qquad x(0) = \mathbf{0}. \qquad (31)$$

(b) State augmentation

Define the ρ_{aug}-dimensional vector $\chi_{aug}(t) := \begin{bmatrix} x(t) \\ \chi_o(t) \end{bmatrix} = \begin{bmatrix} \mathbf{I}_R \\ \mathbf{V_o}^T \mathbf{I_{a,\bullet}} \end{bmatrix} x(t)$ that includes all extents and extent directions, with $\rho_{aug} = R + \rho_o$. The dynamic behavior of $\chi_{aug}(t)$ can be described by a differential-algebraic system including the R differential Equation (31) and the ρ_o algebraic expressions (16):

$$\dot{x}(t) = V f\left(\overline{\chi}_o(t), x(t), \theta\right), \qquad\qquad x(0) = \mathbf{0} \qquad (32)$$
$$\chi_o(t) = \mathbf{V_o}^T \mathbf{I_{a,\bullet}}\, x(t). \qquad\qquad\qquad\qquad (33)$$

(c) Interpolation of the observable extents and extent directions

To increase the efficiency of system partitioning, it is useful to account for the fact that the observable extents and extent directions can be expressed in terms of measurements. However, since the observable extents and extent directions are only known at H discrete measurement points, their values always need to be obtained via interpolation. In this work, we apply piece-wise linear interpolation as follows:

$$\tilde{\chi}_{o,i}(t) := \tilde{\chi}_{o,h} + \frac{t - t_h}{t_{h+1} - t_h}\left(\tilde{\chi}_{o,h+1} - \tilde{\chi}_{o,h}\right), \qquad t_h \le t < t_{h+1}, \qquad h = 1, \ldots, H, \qquad (34)$$

with which the system (32) and (33) becomes:

$$\dot{x}(t) = V f\left(\tilde{\chi}_{o,i}(t), x(t), \theta\right), \qquad\qquad x(0) = \mathbf{0} \qquad (35)$$
$$\chi_o(t) = \mathbf{V_o}^T \mathbf{I_{a,\bullet}}\, x(t). \qquad\qquad\qquad\qquad (36)$$

3.6.2. Step 2—Graph-Based System Partitioning

The equation system (35) and (36) is now analyzed by means of a graph partitioning procedure to determine the smallest groups of kinetic parameters that can be estimated separately. To this end, the following steps are performed:

(a) Create a graph
One creates a directed graph $\mathcal{F}$ with a vertex for every state variable in $\chi_{aug}(t)$ and every parameter in θ. Hence, this graph has $R + \rho_o + T$ vertices. A directed arc is added from vertex v to vertex w if the vth element of $\begin{bmatrix} \chi_{aug}(t) \\ \theta \end{bmatrix}$ appears in the right-hand side of the wth equation in (35) and (36) ($v = 1, \ldots, R + \rho_o + T$, $w = 1, \ldots, R + \rho_o$). This graph represents the information flow for simulating (35) and (36). Additional arcs and vertices may be added to describe the influence of known inputs and the links between extents and measured variables. For system partitioning, this is however unnecessary and omitted for clarity.

(b) Extents predicted from measurements or simulation
The simultaneous approach uses a complete model of the reaction system to predict the extents (or concentrations) via simulation. If one wants to partition the reaction system into small groups of reactions, only the extents belonging to a given group can be generated via the simulation of that group. The other extents that enter the rate laws must be provided by the user as quantities known from measurements.

That information can be included in the graph $\mathcal{F}$ by annotating the various arcs. The arcs that originate at a vertex corresponding to an observable extent or an observable extent direction are labeled *observation arcs*, considering that observable extent or an observable extent direction can be replaced with their measured values (34). They are visualized as dashed-line arrows. The remaining arcs are labeled *simulation arcs* and visualized as solid-line arrows. The observation arcs represent the idea that the elements of $\tilde{\tilde{\chi}}_{\mathrm{o,i}}(t)$ can be regarded as known inputs for simulating (35) and (36).

(c)   *Subgraph selection*

Identify the J subgraphs $\mathcal{F}^{(j)}$ consisting of arcs and vertices in $\mathcal{F}$ on directed paths that (i) lead to a vertex representing an observable extent or an observable extent direction; and (ii) consist of simulation arcs only. The selected vertices represent an $R^{(j)}$-dimensional vector of extents $x^{(j)}(t)$, a $\rho_{\mathrm{o}}^{(j)}$-dimensional vector of directions $\chi_{\mathrm{o}}^{(j)}(t)$, and a $T^{(j)}$-dimensional vector of parameters $\theta^{(j)}$. The positions of $x^{(j)}(t)$ in $x(t)$ are given by the vector $\mathbf{j}$ so that:

$$x^{(j)}(t) := \mathbf{I}_{\mathbf{j},\bullet} \, x(t) \tag{37}$$

$$f^{(j)}(\bullet) := \mathbf{I}_{\mathbf{j},\bullet} \, f(\bullet), \tag{38}$$

and the selection matrices $\mathbf{\Lambda}_{\mathrm{o}}^{(j)}$ and $\mathbf{\Lambda}_{\theta}^{(j)}$ are defined so that:

$$\chi_{\mathrm{o}}^{(j)}(t) := \mathbf{\Lambda}_{\mathrm{o}}^{(j)} \, \chi_{\mathrm{o}}(t) \tag{39}$$

$$\theta^{(j)} := \mathbf{\Lambda}_{\theta}^{(j)} \, \theta. \tag{40}$$

This means that each subgraph $\mathcal{F}^{(j)}$ represents a subset of Equations (35) and (36) that describes the dynamics of $x^{(j)}(t)$ and $\chi_{\mathrm{o}}^{(j)}(t)$ without reference to any other state variable:

$$\dot{x}^{(j)}(t) = V \, f^{(j)}\left(\tilde{\tilde{\chi}}_{\mathrm{o,i}}(t), x^{(j)}(t), \theta^{(j)}\right), \qquad\qquad x^{(j)}(0) = x_{\mathrm{o}} \tag{41}$$

$$\chi_{\mathrm{o}}^{(j)}(t) = \mathbf{U}^{(j)} \, x^{(j)}(t), \tag{42}$$

with $\mathbf{U}^{(j)} := \mathbf{\Lambda}_{\mathrm{o}}^{(j)} \, \mathbf{V}_{\mathrm{o}}{}^{\mathsf{T}} \, \mathbf{I}_{\mathbf{a},\bullet} \, \mathbf{I}_{\mathbf{j},\bullet}{}^{\mathsf{T}}$.

(d)   *Add observation arcs and vertices*

For every graph $\mathcal{F}^{(j)}$, add (i) the observation arcs that have a target vertex belonging to $\mathcal{F}^{(j)}$ and (ii) the source vertices of the added observation arcs. These added source vertices represent the minimal subset of interpolants in $\tilde{\tilde{\chi}}_{\mathrm{o,i}}(t)$ (34) that are required to simulate $x^{(j)}(t)$ and $\chi_{\mathrm{o}}^{(j)}(t)$ and are referred to as $\tilde{\tilde{\chi}}_{\mathrm{o,i}}^{(j)}(t)$. This means that the graph $\mathcal{F}^{(j)}$ now represents all information required to simulate the observable extents $x^{(j)}(t)$ and the observable extent directions $\chi_{\mathrm{o}}^{(j)}(t)$. Accordingly, one can rewrite the jth equation subsystem as:

$$\dot{x}^{(j)}(t) = V \, f^{(j)}\left(\tilde{\tilde{\chi}}_{\mathrm{o,i}}^{(j)}(t), \chi_{\mathrm{o}}^{(j)}(t), x^{(j)}(t), \theta^{(j)}\right), \qquad\qquad x^{(j)}(0) = x_{\mathrm{o}} \tag{43}$$

$$\chi_{\mathrm{o}}^{(j)}(t) = \mathbf{U}^{(j)} \, x^{(j)}(t). \tag{44}$$

At the end of this procedure, the equation system (35) and (36) is approximated by J smaller equation subsystems, each including a subset of the kinetic parameters. As intended, every kinetic parameter appears in at most one of the J subsystems. In addition, the identified subsystems do not

share any of the observable extents or extent directions as state variables, that is, each observable extent and extent direction is simulated in only one of the identified subsystems.

3.7. Parameter Estimation Methods

In this work, we solve the parameter estimation problem in two distinct ways. The first way consists in a simultaneous estimation of all parameters in the maximum-likelihood sense. The second way consists in an extent-based incremental parameter estimation. The next paragraphs describe how incremental parameter estimation approximates the simultaneous estimation procedure to minimize the number of parameters that are estimated together.

3.7.1. Simultaneous Parameter Estimation

Maximum-likelihood estimation of the kinetic parameters can be obtained by solving minimization of the weighted mean squared error:

$\underline{\text{Problem } P_0}$

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} Q_0 := \frac{1}{H \cdot M} \sum_{h}^{H} \left(\tilde{\boldsymbol{y}}_h - \boldsymbol{y}(t_h)\right)^{\mathsf{T}} \boldsymbol{\Sigma}_\epsilon^{-1} \left(\tilde{\boldsymbol{y}}_h - \boldsymbol{y}(t_h)\right), \tag{45}$$

subject to (1)–(3). This estimation problem can be equivalently formulated in terms of the computed extents as follows:

$\underline{\text{Problem } P_0^*}$

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} Q_0^* := \sum_{h}^{H} \left(\tilde{\overline{\boldsymbol{\chi}}}_{\mathsf{o},h} - \overline{\boldsymbol{\chi}}_{\mathsf{o}}(t_h)\right)^{\mathsf{T}} \mathbf{W} \left(\tilde{\overline{\boldsymbol{\chi}}}_{\mathsf{o},h} - \overline{\boldsymbol{\chi}}_{\mathsf{o}}(t_h)\right), \tag{46}$$

subject to (6)–(8) and (20) and with $\mathbf{W} := \boldsymbol{\Sigma}_{\overline{\chi}}^{-1}$.

3.7.2. Incremental Parameter Estimation

The incremental parameter estimation procedure is obtained by applying two modifications, A and B, to problem P_0^*.

(a) *Modification A: Removing correlation terms.* Let $A^{(j)} = R_{\mathsf{o}}^{(j)} + \rho_{\mathsf{o}}^{(j)}$ and define the $A^{(j)}$-dimensional vector of observable extents and extent directions $\overline{\boldsymbol{\chi}}_{\mathsf{o}}^{(j)}(t)$. This vector includes all observable extents $\boldsymbol{x}_{\mathsf{o}}^{(j)}(t)$, whose positions in $\boldsymbol{x}^{(j)}(t)$ are given by the vector $\mathbf{o}^{(j)}$, and all observable extent directions $\boldsymbol{\chi}_{\mathsf{o}}^{(j)}(t)$ in Subsystem j. We further define the matrix $\overline{\mathbf{U}}^{(j)}$ so that:

$$\boldsymbol{x}_{\mathsf{o}}^{(j)}(t) := \mathbf{I}_{\mathbf{o}^{(j)},\bullet}\, \boldsymbol{x}^{(j)}(t) \tag{47}$$

$$\overline{\boldsymbol{\chi}}_{\mathsf{o}}^{(j)}(t) = \begin{bmatrix} \boldsymbol{x}_{\mathsf{o}}^{(j)}(t) \\ \boldsymbol{\chi}_{\mathsf{o}}^{(j)}(t) \end{bmatrix} = \overline{\mathbf{U}}^{(j)}\, \boldsymbol{x}^{(j)}(t) \tag{48}$$

$$\overline{\mathbf{U}}^{(j)} := \begin{bmatrix} \mathbf{I}_{\mathbf{o}^{(j)},\bullet} \\ \mathbf{U}^{(j)} \end{bmatrix}. \tag{49}$$

The $A^{(j)}$-dimensional vector $\tilde{\overline{\boldsymbol{\chi}}}_{\mathsf{o},h}^{(j)}$ is then defined by selecting the elements of $\tilde{\overline{\boldsymbol{\chi}}}_{\mathsf{o},h}$ as:

$$\tilde{\overline{\boldsymbol{\chi}}}_{\mathsf{o},h}^{(j)} := \mathbf{I}_{\mathbf{e}^{(j)},\bullet}\, \tilde{\overline{\boldsymbol{\chi}}}_{\mathsf{o},h}, \tag{50}$$

where the vector $\mathbf{e}^{(j)}$ gives the positions of $\tilde{\tilde{\chi}}_{\mathrm{o},h}^{(j)}$ in $\tilde{\tilde{\chi}}_{\mathrm{o},h}$. Now define the residuals $d_h^{(j)}$ associated with subsystem j:

$$d_h^{(j)} := \tilde{\tilde{\chi}}_{\mathrm{o},h}^{(j)} - \overline{\chi}_{\mathrm{o}}^{(j)}(t_h) \tag{51}$$

This way, the objective function Q_0^* defined above can be reformulated as:

$$Q_0^* = \sum_h^H \sum_i^J \sum_j^J d_h^{(i)\,\mathrm{T}} \mathbf{I}_{\mathbf{e}^{(i)},\bullet} \, \mathbf{W} \, \mathbf{I}_{\mathbf{e}^{(j)},\bullet}^{\mathrm{T}} d_h^{(j)}$$

$$= \sum_h^H \left(\sum_j^J d_h^{(j)\,\mathrm{T}} \mathbf{W}^{(j)} \, d_h^{(j)} + \sum_{i,j:\,i\neq j} d_h^{(i)\,\mathrm{T}} \mathbf{I}_{\mathbf{e}^{(i)},\bullet} \, \mathbf{W} \, \mathbf{I}_{\mathbf{e}^{(j)},\bullet}^{\mathrm{T}} d_h^{(j)} \right) \tag{52}$$

and can subsequently approximated with Q_1:

$$Q_0^* \approx Q_1 := \sum_j^J \sum_h^H d_h^{(j)\,\mathrm{T}} \mathbf{W}^{(j)} \, , d_h^{(j)} \tag{53}$$

with $\mathbf{W}^{(j)} := \mathbf{I}_{\mathbf{e}^{(j)},\bullet} \, \mathbf{W} \, \mathbf{I}_{\mathbf{e}^{(j)},\bullet}^{\mathrm{T}}$. This first modification results in:

Problem P_1

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} Q_1 := \sum_j^J \sum_h^H d_h^{(j)\,\mathrm{T}} \mathbf{W}^{(j)} \, d_h^{(j)}, \tag{54}$$

subject to (6)–(8) and (20). This error due to this approximation is referred to as *Type A* approximation error. This error is zero if all matrices $\mathbf{I}_{\mathbf{e}^{(i)},\bullet} \, \mathbf{W} \, \mathbf{I}_{\mathbf{e}^{(j)},\bullet}^{\mathrm{T}}, i \neq j$, are zero matrices. This is true when the correlation between the estimation errors of any element in $\tilde{\tilde{\chi}}_{\mathrm{o},h}^{(i)}$ and any element of $\tilde{\tilde{\chi}}_{\mathrm{o},h}^{(j)}$ ($i \neq j$) is zero.

(b)   *Modification B: Separation of problem P_1 into J smaller problems $P_1^{(j)}$.* An approximation to problem P_1 is now obtained by optimizing each of the J terms in (54) separately and simulating the values of $\overline{\chi}_{\mathrm{o}}^{(j)}(t_h)$ with (43) and (44). We refer to each problem as $P_1^{(j)}$:

Problem $P_1^{(j)}$

$$\hat{\boldsymbol{\theta}}^{(j)} = \arg\min_{\boldsymbol{\theta}} Q_1^{(j)} := \sum_h^H d_h^{(j)\,\mathrm{T}} \mathbf{W}^{(j)} \, d_h^{(j)} \tag{55}$$

subject to

$$\dot{x}^{(j)}(t) = V \, f^{(j)} \left(\tilde{\tilde{\chi}}_{\mathrm{o,i}}^{(j)}(t), \chi_{\mathrm{o}}^{(j)}(t), x^{(j)}(t), \boldsymbol{\theta}^{(j)} \right), \qquad x^{(j)}(0) = \mathbf{0} \tag{56}$$

$$\chi_{\mathrm{o}}^{(j)}(t) = \mathbf{U}^{(j)} \, x^{(j)}(t) \tag{57}$$

$$d_h^{(j)} = \tilde{\tilde{\chi}}_{\mathrm{o},h}^{(j)} - \overline{\chi}_{\mathrm{o}}^{(j)}(t_h) = \tilde{\tilde{\chi}}_{\mathrm{o},h}^{(j)} - \overline{\mathbf{U}}^{(j)} \, x^{(j)}(t_h). \tag{58}$$

The most important feature of problems $P_1^{(j)}$, $j = 1, \dots, J$, is that each of them involves only a

small number of parameters $\theta^{(j)}$. Please note that the approximation of problem P_1 by this set of problems $P_1^{(j)}$ is perfect in the special case where the right-hand sides (56) do not involve interpolated extents, that is, if the vectors $\tilde{\tilde{\chi}}_{o,i}^{(j)}(t)$ $(j = 1, \ldots, J)$ are empty. Another error, named *Type B* approximation error, is introduced when this is not the case.

Thanks to two modifications of problem P_0, one obtains J smaller problems $P_1^{(j)}$, each of which includes only a fraction of the original set of parameters. The price to pay for such a simplification is a deviation from maximum likelihood due to the introduction of approximation errors (type A & B). These errors are often marginal as demonstrated below.

As in previous work [5–7], the parameter estimates obtained by solving problem $P_1^{(j)}$ $(j = 1, \ldots, J)$ can serve as reliable initial guesses to initiate the solution to problem P_0. The problem P_0 when solved with these initial guesses is named problem P_{1+0}.

3.8. Implementation

All results can be reproduced with the open-source Efficient Model Identification (EMI) MATLAB package for efficient model identification [6] that includes all methods and simulations used in this study. This package is added in the Supplementary Information (Section A).

4. Results

We first explain the results obtained within an extensive simulation study to demonstrate the method. After that, results obtained with an experimental data set are used to demonstrate real-world applicability.

4.1. Simulation Study

The developed methods are demonstrated via a batch reaction system and investigating 5 different measurement scenarios (A–E). The reaction system is described first. Then, scenario A and its results are discussed in detail. The results for scenarios B to E are only summarized, with the details described in the Supplementary Information (Section B).

4.1.1. Reaction System

This reaction system has $R = 5$ reactions involving $S = 6$ species (A to F) with the following reaction scheme:

$$
\begin{aligned}
R_1 : \quad & A + B \longrightarrow C \\
R_2 : \quad & 2\,A \longrightarrow D \\
R_3 : \quad & 2\,C \longrightarrow B + D \\
R_4 : \quad & D \longrightarrow E \\
R_5 : \quad & 2\,D \longrightarrow E + F
\end{aligned}
$$
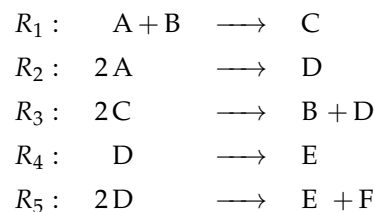
with

$$
\mathbf{N} := \begin{bmatrix}
-1 & -1 & +1 & 0 & 0 & 0 \\
-2 & 0 & 0 & +1 & 0 & 0 \\
0 & +1 & -2 & +1 & 0 & 0 \\
0 & 0 & 0 & -1 & +1 & 0 \\
0 & 0 & 0 & -2 & +1 & +1
\end{bmatrix}.
\tag{59}
$$

4.1.2. Dynamic Model in Terms of Numbers of Moles

The simulated kinetic rate expressions are:

$$
\boldsymbol{f}\left(\boldsymbol{n}(t), \boldsymbol{\theta}\right) := \begin{bmatrix} k_1 \left(\frac{n_1(t)}{V} \frac{n_2(t)}{V} - K_1 \frac{n_3(t)}{V} \right) \\ k_2 \left(\frac{n_1(t)}{V} \right)^2 \\ k_3 \frac{n_3(t)}{V} \\ k_4 \frac{n_4(t)}{V} \\ k_5 \left(\frac{n_4(t)}{V} \right)^2 \end{bmatrix} \tag{60}
$$

with $\boldsymbol{\theta} := \begin{bmatrix} k_1 & k_2 & k_3 & k_4 & k_5 & K_1 \end{bmatrix}^{\mathrm{T}}$. The ground truth values of the kinetic parameters are given in Table 2.

Table 2. Scenario A—Parameter estimates. Model simulation and parameter estimation results: kinetic parameters and model fit. All parameter estimates are reported with their standard deviation based on the Laplacian approximation of the likelihood function (i.e., $exp\left(-Q_0\right)$, $exp\left(-Q_1^{(j)}\right)$).

	Name	Unit	True Value	P_0	$P_1^{(j)}$	P_{1+0}
θ_1	k_1	$\mathrm{L\,mol^{-1}\,h^{-1}}$	2.0	1.9490 ($\pm$0.0003)	2.0010 ($\pm$3.3625)	1.9857 ($\pm$0.0001)
θ_2	k_2	$\mathrm{L\,mol^{-1}\,h^{-1}}$	0.5	0.4929 ($\pm$0.0002)	0.5006 ($\pm$2.4090)	0.4771 ($\pm$0.0001)
θ_3	k_3	$\mathrm{h^{-1}}$	1.0	1.0056 ($\pm$0.0827)	0.8955 ($\pm$0.0002)	0.9974 ($\pm$0.0805)
θ_4	k_4	$\mathrm{h^{-1}}$	0.4	0.4342 ($\pm$0.4275)	0.4003 ($\pm$2.6400)	0.4606 ($\pm$0.1309)
θ_5	k_5	$\mathrm{L\,mol^{-1}\,h^{-1}}$	1.6	1.2285 ($\pm$2.5160)	1.6009 ($\pm$2.9587)	1.2111 ($\pm$0.8018)
θ_6	K_1	$\mathrm{mol\,L^{-1}}$	1.4	1.3435 ($\pm$0.1016)	1.4001 ($\pm$2.9081)	1.3986 ($\pm$0.1135)
$\sqrt{\frac{Q_0}{H\cdot M}}$	WRMSR		$-$	0.44731	0.47934	0.44771

The initial conditions are $\boldsymbol{n}_0 := \begin{bmatrix} 0.73 & 0.42 & 0 & 0 & 0 & 0 \end{bmatrix}^{\mathrm{T}}$ mol, and the reactor volume is $V = 1L$. In all investigated scenarios, measurements are taken at intervals of 5 min during 10 h ($H = 121$, $t_h = 0, 1/12, 2/12, \ldots, 10$).

4.1.3. Scenario A

The first scenario considers the case where the concentrations of B and C and the sum of the concentrations of E and F are measured, that is,

$$
\mathbf{M} := \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}. \tag{61}
$$

The measurement error-covariance matrix is $\boldsymbol{\Sigma}_\epsilon = \mathrm{diag}\left(\begin{bmatrix} 1 & 1 & 2 \end{bmatrix}^{\mathrm{T}} \right) 10^{-4}\,\mathrm{mol^2 L^{-2}}$. Figure 2 shows the simulated noise-free and noisy measurements.

Extent labeling

To specify the measurement model in terms of extents (7), one needs to specify $\boldsymbol{y}_0$ and $\mathbf{G}$:

$$
\boldsymbol{y}_0 = \mathbf{M}\,\boldsymbol{n}_0 = \begin{bmatrix} 0.42 & 0 & 0 \end{bmatrix}^{\mathrm{T}} \tag{62}
$$

$$
\mathbf{G} = \mathbf{M}\,\mathbf{N}^{\mathrm{T}} = \begin{bmatrix} -1 & 0 & +1 & 0 & 0 \\ +1 & 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & +1 & +2 \end{bmatrix}. \tag{63}
$$

The reduced row echelon form of **G** is

$$\mathbf{B} = \begin{bmatrix} +1 & 0 & 0 & 0 & 0 \\ 0 & 0 & +1 & 0 & 0 \\ 0 & 0 & 0 & +1 & +2 \end{bmatrix}, \tag{64}$$

which indicates that there are:

- $R_n = 1$ non-sensed extent, $x_n(t) = x_2(t)$,
- $R_o = 2$ observable extents, $x_o(t) = \begin{bmatrix} x_1(t) & x_3(t) \end{bmatrix}^T$,
- $R_a = 2$ ambiguous extents, $x_a(t) = \begin{bmatrix} x_4(t) & x_5(t) \end{bmatrix}^T$,

which gives:

$$\mathbf{G}_{\bullet,\mathbf{a}} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ +1 & +2 \end{bmatrix}. \tag{65}$$

<u>Observable extents and extent directions</u>

The factorization of $\mathbf{G}_{\bullet,\mathbf{a}}$ gives:

$$\mathcal{G}_o = \begin{bmatrix} 0 \\ 0 \\ +1 \end{bmatrix}, \qquad \mathbf{V}_o^T = \begin{bmatrix} +1 & +2 \end{bmatrix}, \tag{66}$$

so that

$$\chi_o(t) = \mathbf{V}_o^T \mathbf{I}_{\mathbf{a},\bullet} \, x(t) = \begin{bmatrix} 0 & 0 & 0 & 1 & 2 \end{bmatrix} x(t) = x_4(t) + 2\, x_5(t) \tag{67}$$

$$\overline{\chi}_o(t) = \begin{bmatrix} x_1(t) \\ x_3(t) \\ \chi_o(t) \end{bmatrix} = \begin{bmatrix} x_1(t) \\ x_3(t) \\ x_4(t) + 2\, x_5(t) \end{bmatrix}. \tag{68}$$

The values of the observable extents and extent directions $\tilde{\overline{\chi}}_{o,h} := \begin{bmatrix} \tilde{x}_{1,h} & \tilde{x}_{3,h} & \tilde{\chi}_{o,h} \end{bmatrix}^T$ can be computed from (23) with

$$\mathbf{P} = \begin{bmatrix} -2 & -1 & 0 \\ -1 & -1 & 0 \\ 0 & 0 & +1 \end{bmatrix}. \tag{69}$$

The expected estimation error variance-covariance matrix $\Sigma_{\overline{\chi}}$ is:

$$\Sigma_{\overline{\chi}} = \begin{bmatrix} +5 & +3 & 0 \\ +3 & +2 & 0 \\ 0 & 0 & +2 \end{bmatrix} 10^{-4}. \tag{70}$$

This demonstrates that the estimation errors in the first and second observable extents are uncorrelated with the estimation error in the computed observable extent direction. Please note that the correlation between the estimation errors of the first and second observable extents is fairly high ($\frac{3}{\sqrt{5 \cdot 2}} = 0.95$). Figure 3 shows the computed values of the observable extents and extent direction.
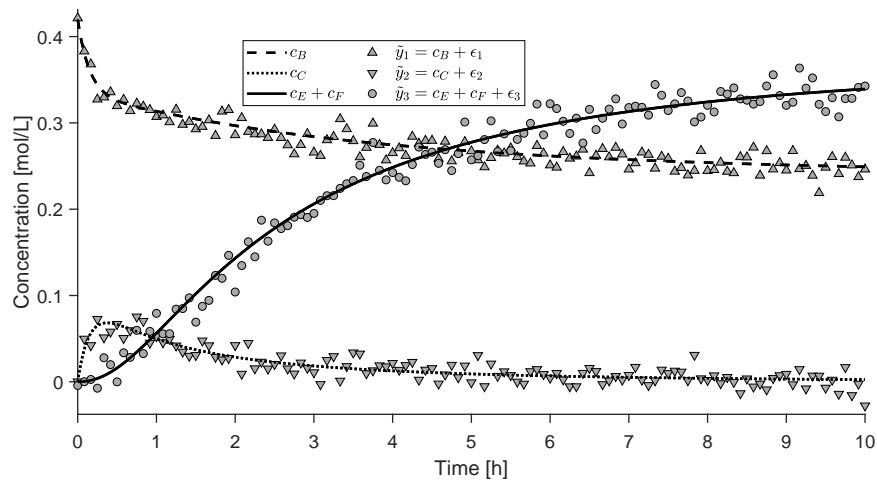
Figure 2. Scenario A—Simulation. Noise-free and noisy measurements as a function of time.
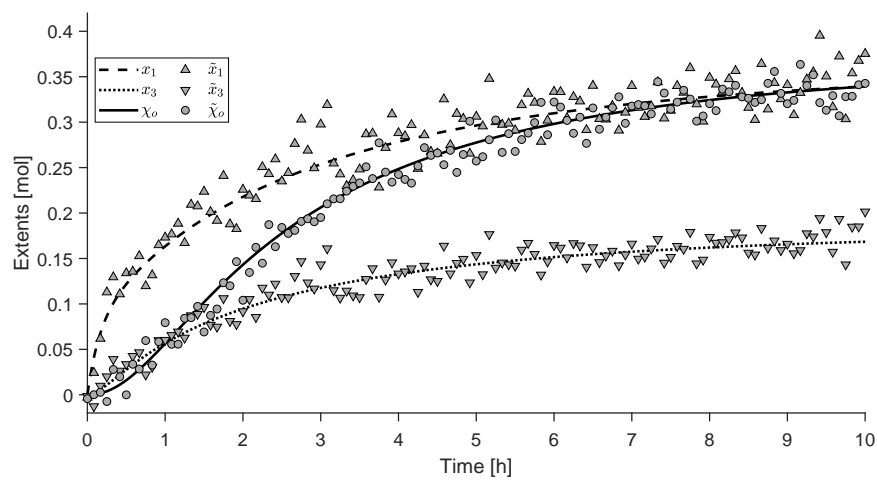
Figure 3. Scenario A—Extent computation. Observable extents (x_1, x_3) and observable extent direction (χ_o) and their computed equivalents ($\tilde{x}_1$, $\tilde{x}_3$, $\tilde{\chi}_o$).

Expression (27) is then fully specified with the following vectors and matrices:

$$
\boldsymbol{n}_0 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}
\tag{71}
$$

$$
\overline{\boldsymbol{\mathcal{N}}}_{\mathsf{o}} := \begin{bmatrix} -1 & -1 & +1 & 0 & 0 & 0 \\ 0 & +1 & -2 & +1 & 0 & 0 \\ 0 & 0 & 0 & -1 & \frac{3}{5} & \frac{2}{5} \end{bmatrix}.
\tag{72}
$$

$$
\overline{\boldsymbol{\mathcal{N}}}_{\mathsf{u}} := \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ -2 & 0 & 0 & +1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & +\frac{2}{5} & -\frac{2}{5} \\ 0 & 0 & 0 & 0 & -\frac{1}{5} & +\frac{1}{5} \end{bmatrix}
\tag{73}
$$

Since the 2nd and 3rd column of $\overline{\mathcal{N}}_{\mathrm{u}}$ are null vectors, it follows that values of $\tilde{n}_2(t)$ and $\tilde{n}_3(t)$ can be computed based on the interpolated estimates $\tilde{\tilde{\chi}}_{\mathrm{o,i}}(t)$ without the need to simulate $x(t)$. A further exploration of this idea remains out of the scope of this paper, however.

System partitioning

For simplicity of notation, we omit the time dependence in what follows. For example, the interpolants $\tilde{\tilde{\chi}}_{\mathrm{o,i}}(t)$ are given as $\tilde{\tilde{\chi}}_{\mathrm{o,i}} := \begin{bmatrix} \tilde{x}_1 & \tilde{x}_3 & \tilde{\chi}_{\mathrm{o}} \end{bmatrix}^{\mathrm{T}}$. Following Steps 1(a)–(b) in Section 3.6.1, the augmented equation system becomes:

$$
\dot{x} = \begin{bmatrix}
\frac{k_1}{V^2} \left[\left(n_{0,1} - \tilde{x}_1 - 2x_2\right) \left(n_{0,2} - \tilde{x}_1 + \tilde{x}_3\right) - K_1 \left(n_{0,3} + \tilde{x}_1 - 2\tilde{x}_3\right) \right] \\
\frac{k_2}{V^2} \left(n_{0,1} - \tilde{x}_1 - 2x_2\right)^2 \\
\frac{k_3}{V} \left(n_{0,3} + \tilde{x}_1 - 2\tilde{x}_3\right) \\
\frac{k_4}{V} \left(n_{0,4} + \tilde{x}_3 - \tilde{\chi}_{\mathrm{o}} + x_2\right) \\
\frac{k_5}{V^2} \left(n_{0,4} + \tilde{x}_3 - \tilde{\chi}_{\mathrm{o}} + x_2\right)^2
\end{bmatrix}, \qquad x\left(0\right) = \mathbf{0} \qquad (74)
$$

$$
\chi_{\mathrm{o}} = \mathbf{V_o}^{\mathrm{T}} \mathbf{I_{a,\bullet}} \, x = \begin{bmatrix} 0 & 0 & 0 & 1 & 2 \end{bmatrix} x = x_4 + 2\,x_5 . \qquad (75)
$$

Figure 4 shows the graph corresponding to the above equation system. The vertices corresponding to the observable extents and extent direction are shaded, while the other vertices are white. The simulation arcs are shown with full-line arrows, while the observation arcs are shown as dashed-line arrows. To identify possible subsystems, one removes all the observation arcs, which results in two subgraphs. The first subgraph includes the parameters k_1, k_2, k_4, k_5, and K_1, which affect the observable quantities x_1 and χ_{o} via a network that also involves the unobservable extents x_2, x_4, and x_5. The second subgraph is much smaller and includes the parameter k_3 that influences the observable extent x_3.
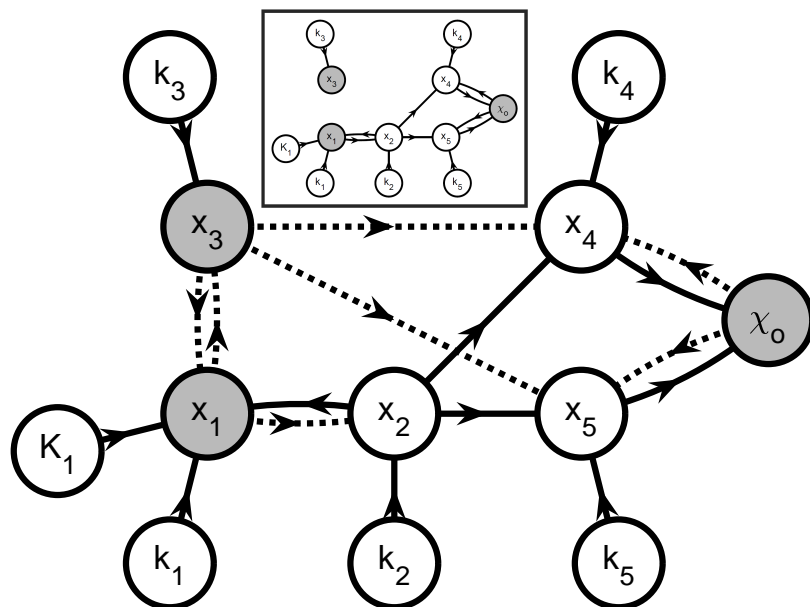
Figure 4. Scenario A—Graph $\mathcal{F}$. There are three shaded vertices corresponding the observable extents (x_1, x_3) and extent direction (χ_{o}). The remaining vertices represent the unobservable extents (x_2, x_4, x_5) and the parameters $(k_1, k_2, k_3, k_4, k_5, K_1)$. The simulation and observation arcs are shown as solid-line and dashed-line arrows, respectively. Removing the observation arcs and graph partitioning results in 2 subgraphs for the parameters as shown in the inset: one graph in which k_1, k_2, k_4, k_5, and K_1 are connected to the observable quantities x_1 and χ_{o}, and another graph in which k_3 is connected to the observable extent x_3.

Accordingly, the $J = 2$ subsystems are:

$$\dot{x}^{(1)} = \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_4 \\ \dot{x}_5 \end{bmatrix} = \begin{bmatrix} \frac{k_1}{V^2} \left[(n_{0,1} - x_1 - 2x_2)(n_{0,2} - x_1 + \tilde{x}_3) \\ - K_1 (n_{0,3} + x_1 - 2\tilde{x}_3) \right] \\ \frac{k_2}{V^2} (n_{0,1} - x_1 - 2x_2)^2 \\ \frac{k_4}{V} (n_{0,4} + \tilde{x}_3 - \chi_o + x_2) \\ \frac{k_5}{V^2} (n_{0,4} + \tilde{x}_3 - \chi_o + x_2)^2 \end{bmatrix}, \quad x^{(1)}(0) = \begin{bmatrix} x_{0,1} \\ x_{0,2} \\ x_{0,4} \\ x_{0,5} \end{bmatrix} = 0 \quad (76)$$

$$\chi_o^{(1)} = \chi_o = x_4 + 2\,x_5, \tag{77}$$

and

$$\dot{x}^{(2)} = \dot{x}_3 = \frac{k_3}{V}(n_{0,3} + \tilde{x}_1 - 2x_3), \qquad x^{(2)}(0) = x_{0,3} = 0. \tag{78}$$

The information flows in these subsystems are shown in the inset of Figure 4.

Incremental parameter estimation

The resulting partitioning means that estimates of the first set of parameters $\theta^{(1)} :=$ $\begin{bmatrix} k_1, k_2, k_4, k_5, K_1 \end{bmatrix}$ can be obtained by minimizing a weighted least-squares deviation between the predicted $x_1(t_h)$ and $\chi_o(t_h)$ and their measured counterparts $\tilde{x}_{1,h}$ and $\tilde{\chi}_{o,h}$, that is, by solving problem $P_1^{(1)}$:

$$\hat{\theta}^{(1)} = \arg\min_{\theta^{(1)}} \sum_h^H d_h^{(1)\,\mathrm{T}} W^{(1)} d_h^{(1)} \tag{79}$$

subject to

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_4(t) \\ \dot{x}_5(t) \end{bmatrix} = \begin{bmatrix} \frac{k_1}{V^2} \left[(n_{0,1} - x_1(t) - 2x_2(t))(n_{0,2} - x_1(t) + \tilde{x}_3(t)) \\ - K_1 (n_{0,3} + x_1(t) - 2\tilde{x}_3(t)) \right] \\ \frac{k_2}{V^2} (n_{0,1} - x_1(t) - 2x_2(t))^2 \\ \frac{k_4}{V} (n_{0,4} + \tilde{x}_3(t) - \chi_o(t) + x_2(t)) \\ \frac{k_5}{V^2} (n_{0,4} + \tilde{x}_3(t) - \chi_o(t) + x_2(t))^2 \end{bmatrix}, \quad \begin{bmatrix} x_1(0) \\ x_2(0) \\ x_4(0) \\ x_5(0) \end{bmatrix} = 0 \quad (80)$$

$$\chi_o(t) = x_4(t) + 2\,x_5(t) \tag{81}$$

$$d_h^{(1)} = \begin{bmatrix} \tilde{x}_{1,h} \\ \tilde{\chi}_{o,h} \end{bmatrix} - \begin{bmatrix} x_1(t_h) \\ \chi_o(t_h) \end{bmatrix}, \tag{82}$$

where

$$W^{(1)} := L_e^{(1)} \Sigma_{\tilde{\chi}}^{-1} L_e^{(1)\,\mathrm{T}} = \begin{bmatrix} +1 & 0 & 0 \\ 0 & 0 & +1 \end{bmatrix} \left(\begin{bmatrix} 2 & -3 & 0 \\ -3 & 5 & 0 \\ 0 & 0 & 1/2 \end{bmatrix} 10^4 \right) \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{\mathrm{T}} \tag{83}$$

$$= \begin{bmatrix} 2 & 0 \\ 0 & 1/2 \end{bmatrix} 10^4. \tag{84}$$

The second set of parameters consists of k_3 only, $\theta^{(2)} := k_3$. Its value can be obtained by minimizing the least-squares deviation between $x_3(t_h)$ and its measured counterparts $\tilde{x}_{3,h}$, that is, by solving problem $P_1^{(2)}$:

$$\hat{\theta}^{(2)} = \arg\min_{\theta^{(2)}} \sum_{h}^{H} \left(d_h^{(2)} \right)^2 \tag{85}$$

subject to

$$\dot{x}_3(t) = \frac{k_3}{V} \left(n_{0,3} + \tilde{x}_1(t) - 2x_3(t) \right), \qquad x_3(0) = 0 \tag{86}$$

$$d_h^{(2)} = \tilde{x}_{3,h} - x_3(t_h). \tag{87}$$

Figure 5 shows the measured concentrations and the simulated profiles obtained with (i) the true parameters; (ii) simultaneous parameter estimation (P_0); and (iii) incremental parameter estimation ($P_1^{(1)}$ and $P_1^{(2)}$). All simulated profiles describe the experiment well and are practically indistinguishable. The parameter values obtained by solving P_0, $P_1^{(1)}$, $P_1^{(2)}$, and P_{1+0} are shown in Table 2. All parameter estimates are in close agreement to each other. It is worth noting that the parameter k_2 can be estimated even if the corresponding extent x_2 is labeled non-sensed, i.e., our results suggest that this parameter is both structurally and practically identifiable.
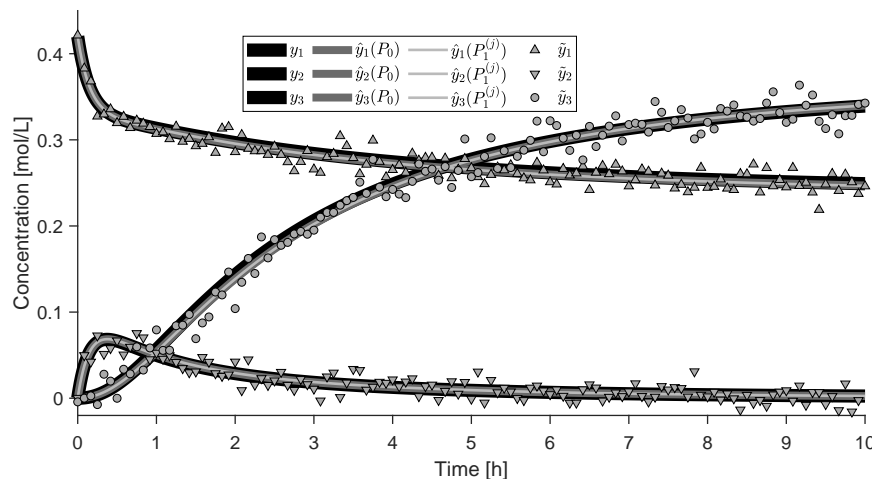
Figure 5. Scenario A—Parameter estimation. Measured concentrations and simulated profiles obtained with (i) true parameters; (ii) simultaneous parameter estimation (P_0); and (iii) incremental parameter estimation ($P_1^{(1)}$ and $P_1^{(2)}$). The produced simulation results exhibit strong overlap, meaning that the identified models closely approximate the ground truth model.

4.1.4. Scenario B

Scenario B assumes concentration measurements for the species B and C only:

$$\mathbf{M} := \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}. \tag{88}$$

The measurement error-covariance matrix is $\mathbf{\Sigma}_\epsilon = \mathrm{diag}\left(\begin{bmatrix} 1 & 1 \end{bmatrix}^{\mathsf{T}} \right) 10^{-4}\,\mathrm{mol^2 L^{-2}}$. The reduced row echelon form of $\mathbf{G}$ is:

$$\mathbf{B} = \begin{bmatrix} +1 & 0 & 0 & 0 & 0 \\ 0 & 0 & +1 & 0 & 0 \end{bmatrix}, \tag{89}$$

which leads to the following extent labeling:

- $R_n = 3$ non-sensed extents, $x_n(t) = \begin{bmatrix} x_2(t) & x_4(t) & x_5(t) \end{bmatrix}^T$,
- $R_o = 2$ observable extents, $x_o(t) = \begin{bmatrix} x_1(t) & x_3(t) \end{bmatrix}^T$,
- $R_a = 0$ ambiguous extents.

Please note that the extent-based method proposed in [8] applies in this case since $R_a = 0$.

Figure 6 shows the corresponding graph $\mathcal{F}$, which can be partitioned into $J = 2$ subsystems and a leftover part. The first subgraph includes the parameters k_1, k_2, and K_1 since they all have some effect on the observable extent x_1 via a network that also involves the unobservable extent x_2. The second subgraph includes the parameter k_3 that influences the observable extent x_3. The information flows in these subsystems are shown in the inset of Figure 6. The leftover part includes the parameters k_4 and k_5 and the extents x_4 and x_5. These parameters and variables are not part of any of the identified subsystems because there are no directed paths composed of simulation arcs from the corresponding vertices to any of the observable extents. Hence, it follows that these parameters are unidentifiable. This is consistent with the method proposed in [8] (not shown). Clearly, the graph-based partitioning procedure shows potential for parameter identifiability analysis, which is discussed as an opportunity below.
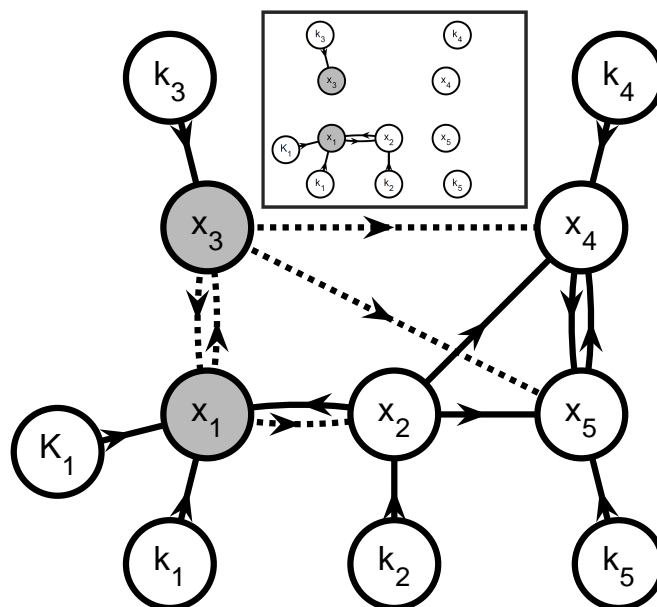
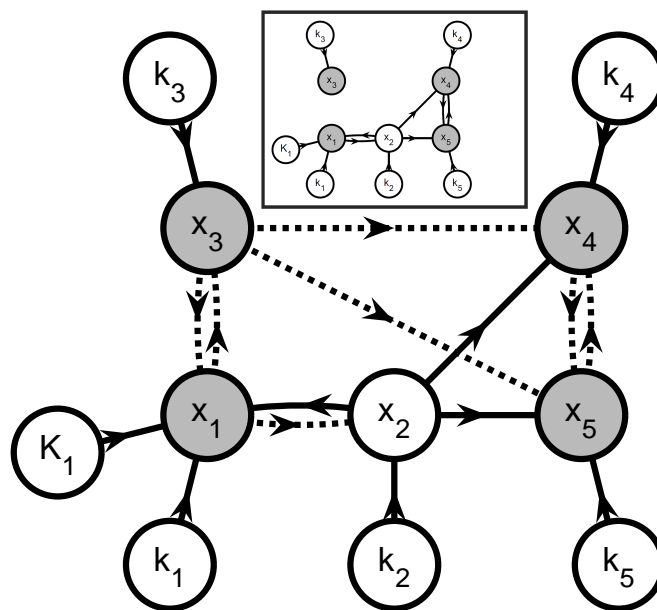**Figure 6.** Scenario B—Graph $\mathcal{F}$. There are two shaded vertices corresponding the observable extents (x_1, x_3). Removing the observation arcs and graph partitioning results in 2 subgraphs for the parameters as shown in the inset: one graph in which k_1, k_2, and K_1 affect the observable extent x_1, a second graph in which k_3 affects the observable extent x_3. The vertices k_4 and k_5 have no effect on the observable extents x_1 and x_3.

4.1.5. Scenario C

Scenario C assumes concentration measurements for the species B, C, E, and F:

$$\mathbf{M} := \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{90}$$

The measurement error-covariance matrix is $\boldsymbol{\Sigma}_\epsilon = \mathrm{diag}\left(\begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}^\mathrm{T}\right) 10^{-4}\,\mathrm{mol}^2\mathrm{L}^{-2}$. The reduced row echelon form of **G** is:

$$\mathbf{B} = \begin{bmatrix} +1 & 0 & 0 & 0 & 0 \\ 0 & 0 & +1 & 0 & 0 \\ 0 & 0 & 0 & +1 & 0 \\ 0 & 0 & 0 & 0 & +1 \end{bmatrix}, \tag{91}$$

which leads to the following extent labeling:

- $R_\mathsf{n} = 1$ non-sensed extents, $\boldsymbol{x}_\mathsf{n}(t) = x_2(t)$.
- $R_\mathsf{o} = 4$ observable extents, $\boldsymbol{x}_\mathsf{o}(t) = \begin{bmatrix} x_1(t) & x_3(t) & x_4(t) & x_6(t) \end{bmatrix}^\mathrm{T}$.
- $R_\mathsf{a} = 0$ ambiguous extents.

Figure 7 shows the corresponding graph $\mathcal{F}$, which can be partitioned into 2 subsystems. The first subgraph includes the parameters k_1, k_2, k_4, k_5, and K_1 since they all have some effect on the observable extents x_1, x_4, and x_5 via a network that also involves the unobservable extents x_2. The second subgraph includes the parameter k_3 that influences the observable extent x_3. The information flows in these subsystems are shown in the inset of Figure 7. Although the assignment of the parameters is the same as in scenario A, the parameter estimation problems are different since the fit of subsystem 1 is determined with respect to $\tilde{x}_1$, $\tilde{x}_4$, and $\tilde{x}_5$ in scenario C, compared to only $\tilde{x}_1$ and $\tilde{\chi}_\mathsf{o}$ in scenario A.

Figure 7. Scenario C—Graph $\mathcal{F}$. There are 4 shaded vertices corresponding to the observable extents x_1, x_3, x_4, and x_5. Removing the observation arcs results in 2 subgraphs for the parameters as shown in the inset: one graph in which k_1, k_2, k_4, k_5, and K_1 affect the observable extents x_1, x_4, and x_5, and another graph in which k_3 affect the observable extent x_3.

4.1.6. Scenario D

Scenario D assumes concentration measurements for the species *A*, *C*, and *E*:

$$\mathbf{M} := \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}. \tag{92}$$

The measurement error-covariance matrix is $\Sigma_\epsilon = \text{diag}\left(\begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^{\mathsf{T}}\right) 10^{-4} \, \text{mol}^2\text{L}^{-2}$. The reduced row echelon form of **G** is:

$$\mathbf{B} = \begin{bmatrix} +1 & 0 & -2 & 0 & 0 \\ 0 & +1 & +1 & 0 & 0 \\ 0 & 0 & 0 & +1 & +1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{93}$$

which leads to the following extent labeling:

- $R_n = 0$ non-sensed extents,
- $R_o = 0$ observable extents,
- $R_a = 5$ ambiguous extents, $x_a(t) = \begin{bmatrix} x_1(t) & x_2(t) & x_3(t) & x_4(t) & x_6(t) \end{bmatrix}^{\mathsf{T}}$.

From $A = \text{rank}(\mathbf{B}) = 3$ and $R_o = 0$, one sees that there are $\rho_o = A - R_o = 3$ observable extent directions among the 5 ambiguous extents. The first two are linear combinations of x_1, x_2, and x_3, while the third one is a linear combination of x_4 and x_5. This particular appearance of extents in subsets of the observable directions stems from the subspace clustering property of the reduced row echelon form and will be discussed in some more detail below.

Figure 8 shows the corresponding graph $\mathcal{F}$, which can be partitioned into 2 subsystems. The first subgraph includes the parameters k_1, k_2, k_3, and K_1 since they all have some effect on the observable extent directions $\chi_{o,1}$ and $\chi_{o,2}$ via a network that also involves the unobservable extents x_1, x_2, and x_3. The second subgraph includes the parameters k_4 and k_5 that influence the observable extent direction $\chi_{o,3}$ via a network that also involves the unobservable extents x_4 and x_5. Hence, this leads to two parameter estimation problems involving 4 and 2 parameters, respectively. The information flows to simulate each of the corresponding subsystems are shown in the inset of Figure 8.
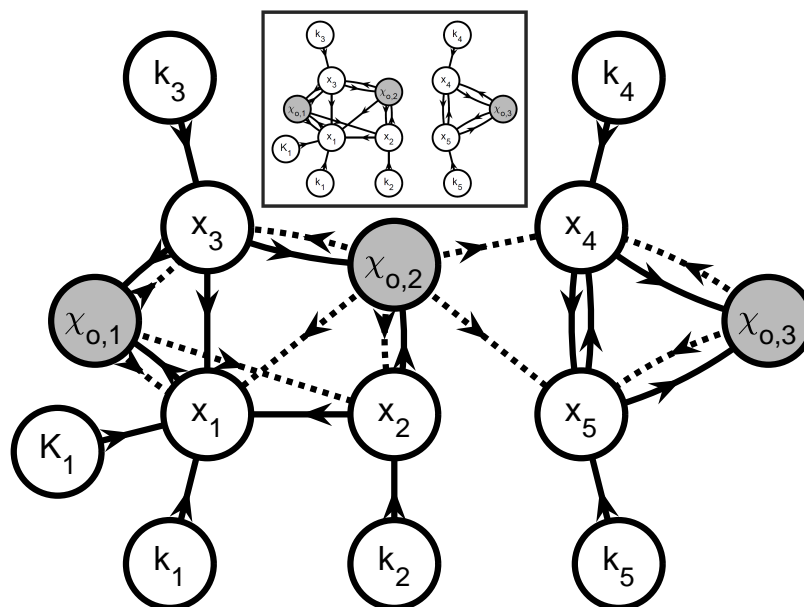
Figure 8. Scenario D—Graph $\mathcal{F}$. There are 3 shaded vertices corresponding the observable extent directions $\chi_{o,1}$, $\chi_{o,2}$, and $\chi_{o,3}$. Removing the observation arcs results in 2 subgraphs for the parameters as shown in the inset: one graph in which k_1, k_2, k_3, and K_1 affect the observable extent directions $\chi_{o,1}$ and $\chi_{o,2}$, and another graph in which k_4 and k_5 affect the observable extent direction $\chi_{o,3}$.

4.1.7. Scenario E

Scenario E assumes concentration measurements of all species:

$$\mathbf{M} := \mathbf{I}_6 \,. \tag{94}$$

The measurement error-covariance matrix is $\boldsymbol{\Sigma}_\epsilon = \mathrm{diag} \left(\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}^\mathrm{T} \right) 10^{-4} \, \mathrm{mol}^2 \mathrm{L}^{-2}$. The reduced row echelon form of $\mathbf{G}$ is:

$$\mathbf{B} = \begin{bmatrix} +1 & 0 & 0 & 0 & 0 \\ 0 & +1 & 0 & 0 & 0 \\ 0 & 0 & +1 & 0 & 0 \\ 0 & 0 & 0 & +1 & 0 \\ 0 & 0 & 0 & 0 & +1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{95}$$

which leads to the following extent labeling:

- $R_\mathrm{n} = 0$ non-sensed extents,
- $R_\mathrm{o} = 5$ observable extents, $\boldsymbol{x_\mathrm{o}}(t) = \begin{bmatrix} x_1(t) & x_2(t) & x_3(t) & x_4(t) & x_6(t) \end{bmatrix}^\mathrm{T}$,
- $R_\mathrm{a} = 0$ ambiguous extents.

This means that the original framework for extent computation, which assumes $A = R$ (Section 3.2.3), applies.

Figure 9 shows the graph $\mathcal{F}$, which can be partitioned into 5 subsystems that include the parameters as follows: (i) k_1 and K_1; (ii) k_2; (iii) k_3; (iv) k_4; and (v) k_5. Hence, one obtains 5 entirely decoupled parameter estimation problems. This set of optimization problems is the same as those obtained with the original extent-based model identification method as is also clear from the inset of Figure 9.
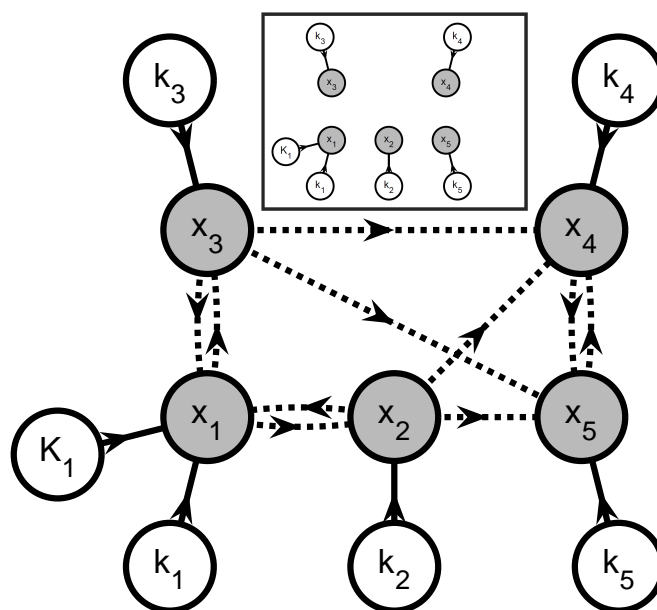
Figure 9. Scenario E—Graph $\mathcal{F}$. All extent vertices are observable and thus shaded. Removing the observation arcs results in 5 subgraphs for the parameters as shown in the inset: k_1 and K_1 are estimated from x_1, k_2 from x_2, k_3 from x_3, k_4 from x_4, and k_5 from x_5.

4.2. Experimental Study

To explore the applicability of the novel incremental parameter estimation method with realistic laboratory data, we execute parameter estimation for the α-pinene batch experiment first reported in [11] and later studied in [12–16]. The reaction network consists of 5 species. These species are labeled A, B, C, D, and E and correspond to α-pinene, dipentene, allo-ocimene, pyronene, and dimer. The reaction network is:

$$
\begin{aligned}
R_1: \quad & A \longrightarrow B \\
R_2: \quad & A \longrightarrow C \\
R_3: \quad & C \longrightarrow D \\
R_4: \quad & C \longrightarrow E \\
R_5: \quad & E \longrightarrow C
\end{aligned}
$$
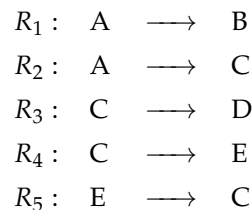
As in [12], the model is defined with the following stoichiometric matrix and process rates:

$$
\mathbf{N} := \begin{bmatrix}
-1 & +1 & 0 & 0 & 0 \\
-1 & 0 & +1 & 0 & 0 \\
0 & 0 & -1 & +1 & 0 \\
0 & 0 & -1 & 0 & +1 \\
0 & 0 & +1 & 0 & -1
\end{bmatrix}
\tag{96}
$$

$$
\boldsymbol{r} := \begin{bmatrix}
k_1 \frac{n_1(t)}{V} \\
k_2 \frac{n_1(t)}{V} \\
k_3 \frac{n_3(t)}{V} \\
k_4 \frac{n_3(t)}{V} \\
k_5 \frac{n_5(t)}{V}
\end{bmatrix}.
\tag{97}
$$

At the start of the batch experiment, only A is present. As explained in [12], the concentrations of all species but D are measured. The concentrations of D are obtained by assuming that 3% of the transformed A is present as D at all times. Finally, all concentrations are normalized to 1 (100%). Figure 10 shows the experimental data reported in [12]. As the initial numbers of moles and the volume are unknown, we express both the concentrations and the extents as dimensionless fractions of the numbers of moles relative to the initial numbers of moles of A. Hence, we can write $\boldsymbol{n}_0 := \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \end{bmatrix}^{\mathrm{T}}$.

Extent labeling

For extent labeling, we assume that all species are measured with independent zero-mean Gaussian measurement errors, similar to the least-squares approach of [12]:

$$
\mathbf{M} := \mathbf{I}_5.
\tag{98}
$$

It follows that

$$
\mathbf{G} = \mathbf{M}\,\mathbf{N}^{\mathrm{T}} = \begin{bmatrix}
-1 & -1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 \\
0 & 1 & -1 & -1 & 1 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & -1
\end{bmatrix}.
\tag{99}
$$

The reduced row echelon form of **G** is

$$\mathbf{B} = \begin{bmatrix} +1 & 0 & 0 & 0 & 0 \\ 0 & +1 & 0 & 0 & 0 \\ 0 & 0 & +1 & 0 & 0 \\ 0 & 0 & 0 & +1 & -1 \end{bmatrix}, \tag{100}$$

which indicates that there are:

- $R_n = 0$ non-sensed extents,
- $R_o = 3$ observable extents, $x_o(t) = \begin{bmatrix} x_1(t) & x_2(t) & x_3(t) \end{bmatrix}^T$,
- $R_a = 2$ ambiguous extents, $x_a(t) = \begin{bmatrix} x_4(t) & x_5(t) \end{bmatrix}^T$,

which delivers:

$$\mathbf{G}_{\bullet,a} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ -1 & +1 \\ 0 & 0 \\ +1 & -1 \end{bmatrix}. \tag{101}$$

Observable extents and extent directions

The factorization of $\mathbf{G}_{\bullet,a}$ gives:

$$\mathcal{G}_o = \begin{bmatrix} 0 \\ 0 \\ 0 \\ +1 \end{bmatrix}, \qquad \mathbf{V}_o^T = \begin{bmatrix} +1 & -1 \end{bmatrix}, \tag{102}$$

so that

$$\chi_o(t) = \mathbf{V}_o^T \mathbf{I}_{a,\bullet} x(t) = \begin{bmatrix} 0 & 0 & 0 & 1 & -1 \end{bmatrix} x(t) = x_4(t) - x_5(t) \tag{103}$$

$$\overline{\chi}_o(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ \chi_o(t) \end{bmatrix} = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) - x_5(t) \end{bmatrix}. \tag{104}$$

Figure 11 shows the computed values of the observable extents and extent direction assuming these definitions. The corresponding error variance-covariance matrix is:

$$\Sigma_{\overline{\chi}} = \begin{bmatrix} +0.8 & -0.6 & -0.2 & -0.2 \\ -0.6 & +1.2 & +0.4 & +0.4 \\ -0.2 & +0.4 & +0.8 & -0.2 \\ -0.2 & +0.4 & -0.2 & +0.8 \end{bmatrix} \sigma \tag{105}$$

with σ the (unknown) measurement error variance.
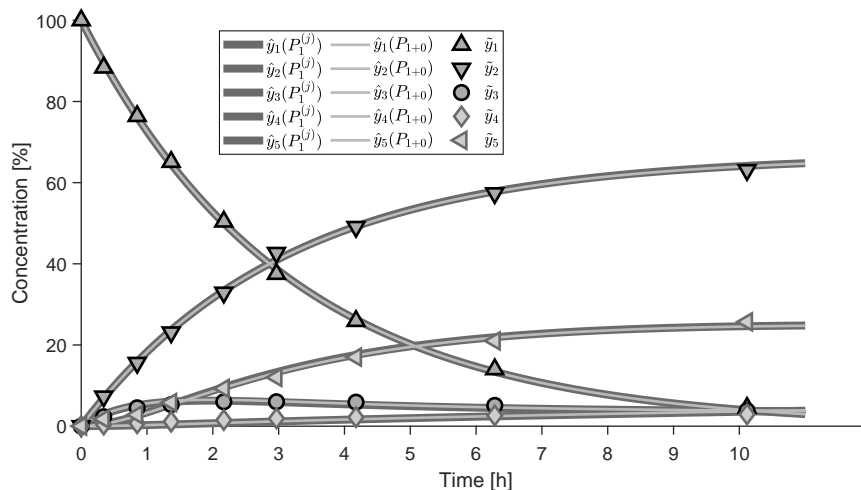
Figure 10. α-pinene—Parameter estimation. Measured concentrations and simulated profiles obtained with (i) incremental parameter estimation ($P_1^{(j)}$, $j = 1, \ldots, 4$) followed by (ii) simultaneous parameter estimation (P_{1+0}). The produced simulation results exhibit strong overlap, meaning that both models produce very similar results.
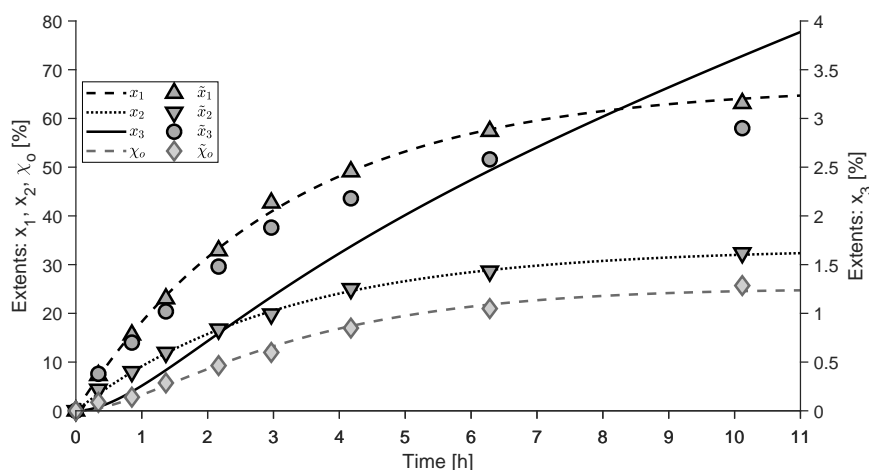
Figure 11. α-pinene—Extents. Computed extents and simulated profiles obtained following incremental parameter estimation ($P_1^{(j)}$, $j = 1, \ldots, 4$).

System partitioning

As above, we omit the time dependence in what follows. Following Steps 1(a)–(b) in Section 3.6.1, the augmented equation system becomes:

$$
\dot{x} = \begin{bmatrix}
k_1 \left(n_{0,1} - \tilde{x}_1 - 2\tilde{x}_2 \right) \\
k_2 \left(n_{0,1} - \tilde{x}_1 - 2\tilde{x}_2 \right) \\
k_3 \left(n_{0,3} + \tilde{x}_2 - \tilde{x}_3 - \tilde{\chi}_\mathrm{o} \right) \\
k_4 \left(n_{0,4} + \tilde{x}_3 \right) \\
k_5 \left(n_{0,5} + \tilde{\chi}_\mathrm{o} \right)
\end{bmatrix}, \qquad x(0) = 0, \tag{106}
$$

$$
\chi_\mathrm{o} = x_4 - x_5. \tag{107}
$$

Figure 12 shows the graph corresponding to the above equation system. Removing all the observation arcs generates 4 subgraphs. The first three subgraphs represent the first three reactions and include: k_1 and x_1 ($j = 1$) ; k_2 and x_2 ($j = 2$); and k_3 and x_3 ($j = 3$). The fourth graph includes

the parameters k_4 and k_5 via a network that also involves the observable extent direction χ_o and the unobservable extents x_4 and x_5 ($j = 4$).
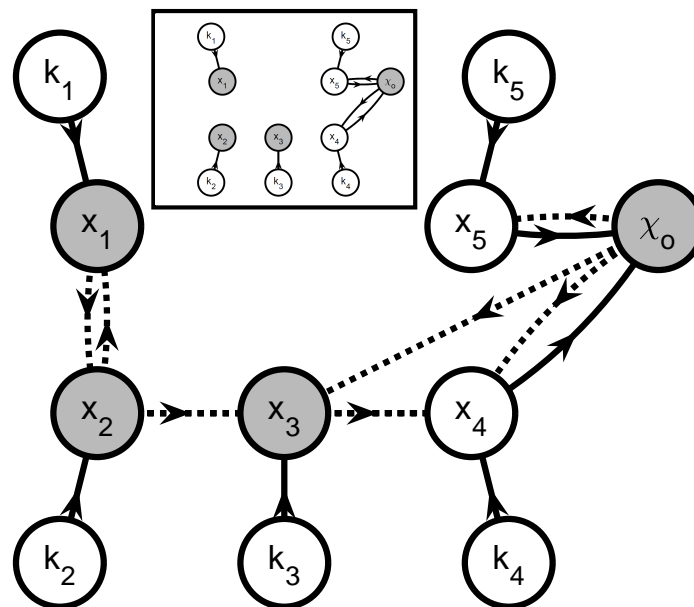
Figure 12. α-pinene—Graph $\mathcal{F}$. There are four shaded vertices corresponding the observable extents (x_1, x_2, x_3) and extent direction (χ_o). The remaining vertices represent the unobservable extents (x_4, x_5) and the parameters (k_1, k_2, k_3, k_4, k_5). The simulation and observation arcs are shown as solid-line and dashed-line arrows, respectively. Removing the observation arcs and graph partitioning results in 4 subgraphs for the parameters as shown in the inset: (i) one with vertices for k_1 and x_1; (ii) one with vertices for k_2 and x_2; (iii) one with vertices for k_3 and x_3; and (iv) one with vertices k_4, k_5, x_4, x_5, and χ_o.

Accordingly, the $J = 4$ subsystems are:

$$\dot{x}^{(1)} = \dot{x}_1 = k_1 \left(n_{0,1} - x_1 - 2\tilde{x}_2\right), \qquad x^{(1)}(0) = x_{0,1} = 0, \tag{108}$$

$$\dot{x}^{(2)} = \dot{x}_2 = k_2 \left(n_{0,1} - \tilde{x}_1 - 2x_2\right), \qquad x^{(2)}(0) = x_{0,2} = 0, \tag{109}$$

$$\dot{x}^{(3)} = \dot{x}_3 = k_3 \left(n_{0,3} + \tilde{x}_2 - x_3 - \tilde{\chi}_\mathrm{o}\right), \qquad x^{(3)}(0) = x_{0,3} = 0, \tag{110}$$

$$\dot{x}^{(4)} = \begin{bmatrix} \dot{x}_4 \\ \dot{x}_5 \end{bmatrix} = \begin{bmatrix} k_4 \left(n_{0,3} + \tilde{x}_2 - \tilde{x}_3 - \tilde{\chi}_\mathrm{o}\right) \\ k_5 \left(n_{0,5} + \tilde{\chi}_\mathrm{o}\right) \end{bmatrix}, \qquad x^{(4)}(0) = \begin{bmatrix} x_{0,4} \\ x_{0,5} \end{bmatrix} = \mathbf{0}, \tag{111}$$

$$\chi_\mathrm{o}^{(4)} = \chi_\mathrm{o} = x_4 - x_5, \tag{112}$$

The information flows in these subsystems are shown in the inset of Figure 12.

Incremental parameter estimation

The resulting partitioning means that estimates of the k_1, k_2, and k_3 can be obtained by solving the following three single-parameter estimation problems ($P_1^{(1)}$, $P_1^{(2)}$, $P_1^{(3)}$):

$$P_1^{(1)}: \quad \hat{k}_1 = \arg\min_{k_1} \sum_h^H \left(\tilde{x}_{1,h} - x_1(t_h)\right)^\mathrm{T} W^{(1)} \left(\tilde{x}_{1,h} - x_1(t_h)\right) \tag{113}$$

subject to

$$\dot{x}^{(1)} = \dot{x}_1 = k_1 \left(n_{0,1} - x_1 - 2\tilde{x}_2\right), \qquad x^{(1)}(0) = x_{0,1} = 0 \tag{114}$$

where

$$W^{(1)} := 5/4 \tag{115}$$

$$P_1^{(2)}: \quad \hat{k}_2 = \arg\min_{k_2} \sum_h^H (\tilde{x}_{2,h} - x_2(t_h))^{\mathrm{T}} W^{(2)} (\tilde{x}_{2,h} - x_2(t_h)) \tag{116}$$

subject to

$$\dot{x}^{(2)} = \dot{x}_2 = k_2 (n_{0,1} - \tilde{x}_1 - 2x_2), \qquad\qquad x^{(2)}(0) = x_{0,2} = 0 \tag{117}$$

where

$$W^{(2)} := 5/3 \tag{118}$$

$$P_1^{(3)}: \quad \hat{k}_3 = \arg\min_{k_3} \sum_h^H (\tilde{x}_{3,h} - x_3(t_h))^{\mathrm{T}} W^{(3)} (\tilde{x}_{3,h} - x_3(t_h)) \tag{119}$$

subject to

$$\dot{x}^{(3)} = \dot{x}_3 = k_3 (n_{0,3} + \tilde{x}_2 - x_3 - \tilde{\chi}_\circ), \qquad\qquad x^{(3)}(0) = x_{0,3} = 0 \tag{120}$$

where

$$W^{(3)} := 5/4 \tag{121}$$

The fourth set of parameters, k_4 and k_5, are estimated by solving the following problem:

$$P_1^{(4)}: \quad \hat{k}_3, \hat{k}_4 = \arg\min_{k_3,k_4} \sum_h^H (\tilde{\chi}_{\circ,h} - \chi_\circ(t_h))^{\mathrm{T}} W^{(4)} (\tilde{\chi}_{\circ,h} - \chi_\circ(t_h)) \tag{122}$$

subject to

$$\dot{x}^{(4)} = \begin{bmatrix} \dot{x}_4 \\ \dot{x}_5 \end{bmatrix} = \begin{bmatrix} k_4 (n_{0,3} + \tilde{x}_2 - \tilde{x}_3 - \tilde{\chi}_\circ) \\ k_5 (n_{0,5} + \tilde{\chi}_\circ) \end{bmatrix}, \qquad\qquad x^{(4)}(0) = \begin{bmatrix} x_{0,4} \\ x_{0,5} \end{bmatrix} = 0 \tag{123}$$

$$\chi_\circ^{(4)} = \chi_\circ = x_4 - x_5, \tag{124}$$

where

$$W^{(4)} := 5/4 \tag{125}$$

Please note that the (unknown) measurement error variance σ can arbitrarily be set equal to one during parameter estimation without affecting the parameter estimates. This, however, means that parameter uncertainties cannot be quantified in this case without estimating σ as well. This estimation and subsequent uncertainty analysis is considered out of scope for this study. Figure 11 shows the simulated extent profiles obtained after solving the optimization problems ($P_1^{(j)}, j = 1, \ldots, 4$). There is a reasonable fit in all cases, except for the third extent of reaction (x_3). This, in the mind of the authors, demonstrates a tangible benefit of the extent-based model framework. Indeed, one can

identify which parts of the model could be improved, e.g., by reconsidering the reaction rate expression corresponding to the computed extents that are approximated poorly. However, this is not explored further in this work.

Figure 10 shows the measured concentrations and the simulated profiles obtained with (i) incremental parameter estimation ($P_1^{(j)}, j = 1, \ldots, 4$), followed by (ii) simultaneous parameter estimation with estimates obtained through incremental parameter estimation (P_{1+0}). The parameter values obtained by solving P_0, $P_1^{(1)}$, $P_1^{(2)}$, and P_{1+0} are shown in Table 3. Note the final values obtained with both P_0 and P_{1+0} are equal within numerical precision.

Table 3. α-pinene—Parameter estimates. The parameters obtained with simultaneous estimation (P_0) and incremental estimation (P_{1+0}) are practically the same.

	Name	Unit	P_0	$P_1^{(j)}$	P_{1+0}
θ_1	k_1	$\%\,\mathrm{h}^{-1}$	0.213	0.214	0.213
θ_2	k_2	$\%\,\mathrm{h}^{-1}$	0.107	0.106	0.107
θ_3	k_3	$\%\,\mathrm{h}^{-1}$	0.074	0.074	0.074
θ_4	k_4	$\%\,\mathrm{h}^{-1}$	0.989	1.037	0.989
θ_5	k_5	$\%\,\mathrm{h}^{-1}$	0.144	0.148	0.144
$\sqrt{\frac{Q_0}{H \cdot M}}$	WRMSR	—	0.66	0.67	0.66

5. Discussion

A novel procedure for model parameter estimation is proposed. The procedure combines two important features: (i) a new framework for extent computation based on the computation of observable directions among the set of sensed but ambiguous extents; and (ii) a graph-based system partitioning procedure that identifies the groups of equations and kinetic parameters that can be simulated independently from the remaining part of the dynamic model. The latter is possible by approximating the original equation system via interpolation of observed extents and extent directions. The proposed procedure enables splitting the parameter estimation problem into smaller problems in cases that could so far only be handled with simultaneous model identification methods (e.g., scenario A, C, and D in Section 4.1). It also subsumes preexisting methods under special conditions such as complete observability (e.g., scenario E in Section 4.1, [5–7]) or absence of ambiguity (e.g., scenario B in Section 4.1, [8]). Thus, this procedure is applicable to parameter estimation for any multivariate differential equation model under a large range of structural observability conditions.

This work removes an important bottleneck for the implementation of incremental model identification to biological systems. Indeed, by providing a method that can be applied to any known stoichiometric matrix $\mathbf{N}$ and measurement matrix $\mathbf{M}$, incremental model identification is now applicable to biological processes where the number of measurements M is lower than the number of modeled reactions R. Computationally speaking, the extent computation and graph-based partitioning are expected to scale well with the number of modeled reactions. The effects on the computational efficiency of the parameter estimation step will however depend primarily on the lengths of the state vectors in the identified subsystems. In turn, these lengths depend strongly on model structure and the available measured variables and less so on the number of states in the complete system.

A remaining obstacle is the fact that both $\mathbf{M}$ and $\mathbf{N}$ are assumed to be known. While several techniques exist to handle this, including methods based on extents [17], some coefficients in these matrices may be unidentifiable based on atomic and stoichiometric balances alone. For this reason, dealing with unknown stoichiometry ($\mathbf{M}$) or ill-defined measurements ($\mathbf{N}$) deserves continued attention.

Another aspect open for exploration is that no global optimization method has been tested so far to solve the smaller parameter estimation problems $P_1^{(j)}$ that result from system partitioning. Inspiration can be drawn from existing parameter estimation methods even if these methods (i) might be complex

yet without publicly available implementation [18]; or (ii) remain limited to cases where all extents are observable and therefore modeled as univariate processes for parameter estimation [19,20].

5.1. Generalized Framework for Extent Computation

The addition of new concepts, such as ambiguous extents and observable extent directions expands the framework for extent computation beyond its original range of applicability. In this study, the focus has been given to the problem of parameter estimation. However, the same generalized framework is likely to be applicable and useful for challenges that have been handled with earlier extent-based methods such as data reconciliation [7], model structure identification [6], and process control [21]. Moreover, there are no obvious limitations that hinder applications involving multi-phase or distributed system models [22,23] or models including algebraic constraints [6,22].

5.2. Optimal System Partitioning

According to our understanding, the proposed procedure delivers the most efficient system partitioning given the available measurements and model information $\mathbf{M}$, $f(\cdot)$, $\mathbf{N}$, and $\boldsymbol{n}_0$. In other words, the procedure generates the subsystems with the smallest number of parameters that can be estimated independently from parameters in other subsystems. However, this is stated here without formal proof. The factorization of $\mathbf{G}_{\bullet,\mathbf{a}}$ based on the reduced row echelon form of $\mathbf{G}$, as described in Section 3.3.1, plays a crucial role. Indeed, it follows from work on subspace clustering methods for noise-free data [24,25] that this factorization minimizes the presence of simulation arcs from the ambiguous extents to the observable directions, thereby leading to the best possible system partitioning.

Please note that the obtained generalized extent framework is optimal only for the purpose of parameter estimation. For instance, one could consider a factorization of $\mathbf{G}_{\bullet,\mathbf{a}}$ which identifies directions that exhibit uncorrelated estimation errors, thereby improving the statistical quality of the computed observable directions and possibly avoiding Type A approximation error. This could be achieved via singular value decomposition of $\mathbf{G}_{\bullet,\mathbf{a}}$. However, singular value decomposition cannot guarantee optimal system partitioning [26]. Consequently, it follows that the optimal factorization of $\mathbf{G}_{\bullet,\mathbf{a}}$ depends on the objective of this factorization. This stands in contrast to the existing body of work concerning extents, which do not involve any such purpose-dependent factorization steps.

Note also that the proposed procedure is designed for the definition of optimality used here. For instance, if one relaxes the constraint that parameters can only appear in one subsystem (Section 3.6), then the proposed graph partitioning procedure is not optimal. For example, in scenario A, one may be able to estimate the parameters in 3 subsets rather than 2, namely, (i) k_1, k_2, and K_1 by simulation of $x_1(t)$ and $x_2(t)$; (ii) k_2, k_3, k_4, and k_5 by simulation of $x_2(t)$, $x_4(t)$, $x_5(t)$, and $\chi_{o,1}(t)$; and (iii) k_3 by simulation of $x_3(t)$. This leads to 3 parameter estimation problems with 3, 4, and 1 parameters, which may be easier to solve than the 2 parameter estimation problems with 5 and 1 parameters as obtained with the proposed procedure. It is however non-trivial to identify a procedure, including both factorization of $\mathbf{G}_{\bullet,\mathbf{a}}$ and system partitioning, that guarantees the best partitioning when applying this relaxed definition of optimality, nor is it clear whether such a procedure exists.

5.3. New Opportunities

This work also hints at several new applications of the generalized framework for extent computation:

(a) **Identifiability analysis.** While not a core objective of this work, it has been shown that the graph partitioning method can help identify unidentifiable parameters. Given that this labeling is based on a model reformulation and does not depend on the temporal resolution or quality of the collected measurements, it follows that the method identifies structurally unidentifiable parameters. Unlike other methods [27], the proposed approach does not require symbolic differentiation. It remains to be explored whether this can also be used to positively identify

structurally identifiable parameters. For a discussion on the evaluation and use of indicators of structural and practical parameter identifiability we refer to [28].

(b) **Soft-sensing.** The appearance of observable directions among the ambiguous extents is closely related to the concepts of observability and detectability in the context of state estimation [29]. Note however that the observability labels in this work are based on the stoichiometric balances and measurement equations alone, thus excluding the dynamic model. At the same time, it is suspected that the extents corresponding to vertices that are not on directed paths to vertices representing observable extents or extent directions can be labeled as structurally unobservable, again observing that the exact timing and quality of the measurements does not play any role in this labeling. Similarly to the identifiability analysis discussed above, such an approach would not rely on symbolic differentiation. Whether this can be used to unambiguously determine observability and detectability for all model states remains to be studied.

(c) **Experimental design.** The labeling of extents and directions as observable or unobservable suggests that experimental design may be used to optimize the selection of measured variables. A method to do so has been applied in [30] to enumerate all Pareto-optimal flow sensor layouts in wastewater treatment plants. In [31], symbolic computation enabled the identification of optimal experimental designs. Similar approaches could be applied as a measurement selection method for metabolic flux analysis and the monitoring of complex systems.

6. Conclusions

In this work, an incremental parameter identification procedure has been developed and tested based on a generalized extent-based framework. This generalized framework enables incremental parameter estimation in cases where the previous methods based on the computation of extents did not permit this. Importantly, through study of simulated and experimental medium-sized examples, the generality of the developed procedure has been demonstrated and new opportunities offered by this framework have been identified.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Rieger, L.; Gillot, S.; Langergraber, G.; Ohtsuki, T.; Shaw, A.; Takács, I.; Winkler, S. *Guidelines for Using Activated Sludge Models. IWA Task Group on Good Modelling Practice. IWA Scientific and Technical Report*; IWA Publishing: London, UK, 2012.
2. Jakeman, A.J.; Letcher, R.A.; Norton, J.P. Ten iterative steps in development and evaluation of environmental models. *Environ. Model. Softw.* **2006**, *21*, 602–614.
3. Bhatt, N.; Amrhein, M.; Bonvin, D. Incremental identification of reaction and mass-transfer kinetics using the concept of extents. *Ind. Eng. Chem. Res.* **2011**, *50*, 12960–12974.
4. Bhatt, N.; Kerimoglu, N.; Amrhein, M.; Marquardt, W.; Bonvin, D. Incremental identification of reaction systems—A comparison between rate-based and extent-based approaches. *Chem. Eng. Sci.* **2012**, *83*, 24–38.
5. Rodrigues, D.; Srinivasan, S.; Billeter, J.; Bonvin, D. Variant and invariant states for chemical reaction systems. *Comput. Chem. Eng.* **2015**, *73*, 23–33.
6. Mašić, A.; Srinivasan, S.; Billeter, J.; Bonvin, D.; Villez, K. Identification of biokinetic models using the concept of extents. *Environ. Sci. Technol.* **2017**, *51*, 7520–7531.
7. Srinivasan, S.; Billeter, J.; Narasimhan, S.; Bonvin, D. Data reconciliation for chemical reaction systems using vessel extents and shape constraints. *Comput. Chem. Eng.* **2017**, *101*, 44–58.
8. Mašić, A.; Billeter, J.; Bonvin, D.; Villez, K. Extent computation under rank-deficient conditions. *IFAC-PapersOnLine* **2017**, *50*, 3929–3934.
9. Kretsovalis, A.; Mah, R.S.H. Observability and redundancy classification in multicomponent process networks. *AIChE J.* **1987**, *33*, 70–82.
10. Crowe, C.M. Observability and redundancy of process data for steady state reconciliation. *Chem. Eng. Sci.* **1989**, *44*, 2909–2917.
11. Fuguitt, R.E.; Hawkins, J.E. Rate of the thermal isomerization of α-Pinene in the liquid phase1. *J. Am. Chem. Soc.* **1947**, *69*, 319–322.
12. Box, G.E.P.; Hunter, W.G.; MacGregor, J.F.; Erjavec, J. Some problems associated with the analysis of multiresponse data. *Technometrics* **1973**, *15*, 33–51.
13. Tjoa, I.B.; Biegler, L.T. Simultaneous solution and optimization strategies for parameter estimation of differential-algebraic equation systems. *Ind. Eng. Chem. Res.* **1991**, *30*, 376–385.
14. Rodriguez-Fernandez, M.; Egea, J.A.; Banga, J.R. Novel metaheuristic for parameter estimation in nonlinear dynamic biological systems. *BMC Bioinform.* **2006**, *2006*, 483.
15. Brunel, N.J.; Clairon, Q. A tracking approach to parameter estimation in linear ordinary differential equations. *Electr. J. Stat.* **2015**, *9*, 2903–2949.
16. Dattner, I.; Gugushvili, S. Application of one-step method to parameter estimation in ODE models. *Stat. Neerl.* **2018**, *72*, 126–156.
17. Bonvin, D.; Rippin, D.W.T. Target factor analysis for the identification of stoichiometric models. *Chem. Eng. Sci.* **1990**, *45*, 3417–3426.
18. Sahlodin, A.M.; Chachuat, B. Convex/concave relaxations of parametric ODEs using Taylor models. *Comput. Chem. Eng.* **2011**, *35*, 844–857.
19. Mašić, A.; Udert, K.; Villez, K. Global parameter optimization for biokinetic modeling of simple batch experiments. *Environ. Model. Softw.* **2016**, *85*, 356–373.
20. Rodrigues, D.; Billeter, J.; Bonvin, D. Maximum-likelihood estimation of kinetic parameters via the extent-based incremental approach. *Comput. Chem. Eng.* **2018**, doi:10.1016/j.compchemeng.2018.05.024.
21. Billeter, J.; Rodrigues, D.; Srinivasan, S.; Amrhein, M.; Bonvin, D. On decoupling rate processes in chemical reaction systems—Methods and applications. *Comput. Chem. Eng.* **2017**, *114*, 296–305.
22. Srinivasan, S.; Billeter, J.; Bonvin, D. Identification of multiphase reaction systems with instantaneous equilibria. *Ind. Eng. Chem. Res.* **2016**, *29*, 8034–8045.
23. Rodrigues, D.; Billeter, J.; Bonvin, D. Generalization of the concept of extents to distributed reaction systems. *Chem. Eng. Sci.* **2017**, *171*, 558–575.
24. Aldroubi, A.; Sekmen, A. Reduced row echelon form and non-linear approximation for subspace segmentation and high-dimensional data clustering. *Appl. Comput. Harmon. Anal.* **2014**, *37*, 271–287.

25. Vidal, R. Subspace clustering. *IEEE Signal Process. Mag.* **2011**, *28*, 52–68.

26. Billeter, J.; Bonvin, D.; Villez, K. *Extent-Based Model Identication under Incomplete Observability Conditions*; Technical Report No. 6, v3.0; Eawag: Dübendorf, Switzerland, 2018.

27. Petersen, B.; Gernaey, K.; Devisscher, M.; Dochain, D.; Vanrolleghem, P.A. A simplified method to assess structurally identifiable parameters in Monod-based activated sludge models. *Water Res.* **2003**, *37*, 2893–2904.

28. Bonvin, D.; Georgakis, C.; Pantelides, C.C.; Barolo, M.; Grover, M.A.; Rodrigues, D.; Schneider, R.; Dochain, D. Linking models and experiments. *Ind. Eng. Chem. Res.* **2016**, *55*, 6891–6903.

29. Sontag, E.D. *Mathematical Control Theory: Deterministic Finite Dimensional Systems*; Springer Science & Business Media: Berlin, Germany, 2013; Volume 6.

30. Villez, K.; Vanrolleghem, P.A.; Corominas, L. Optimal flow sensor placement on wastewater treatment plants. *Water Res.* **2016**, *101*, 75–83.

31. Billeter, J.; Neuhold, Y.M.; Hungerbuehler, K. Systematic prediction of linear dependencies in the concentration profiles and implications on the kinetic hard-modelling of spectroscopic data. *Chemom. Intell. Lab. Syst.* **2009**, *95*, 170–187.