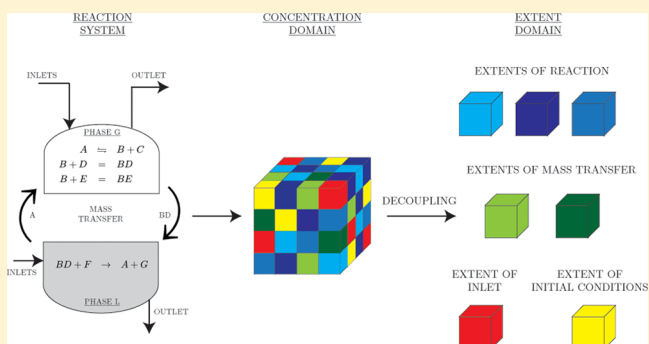


Identification of Multiphase Reaction Systems with Instantaneous Equilibria

Sriniketh Srinivasan, Julien Billeter, and Dominique Bonvin*

Laboratoire d'Automatique, EPFL CH-1015, Lausanne, Switzerland

ABSTRACT: The identification of kinetic models for multiphase reaction systems is complex due to the simultaneous effect of chemical reactions and mass transfers. The extent-based incremental approach simplifies the modeling task by transforming the reaction system into variant states called vessel extents, one for each rate process. This transformation is carried out from the measured numbers of moles (or concentrations) and requires as many measured species as there are rate processes. Then, each vessel extent can be modeled individually, that is, independently of the other dynamic effects. This paper presents a modified version of the extent-based incremental approach that can be used to identify multiphase reaction systems in the presence of instantaneous equilibria. Different routes are possible depending on the number and type of measured species. The approach is illustrated via the simulated example of the oxidation of benzyl alcohol by hypochlorite in a batch reactor.



1. INTRODUCTION

Multiphase chemical reaction systems are often used in the chemical and biochemical industry to convert feed materials into value-added products.¹ The chemical reactions involved in these systems can be either kinetically controlled or nearly instantaneous.^{2,3} Kinetic models are important for process design, scale-up, and operation. Since rapid equilibria affect the concentrations of numerous species, their effects need to be accounted for when developing kinetic models. Note that the complexity of the modeling task is further increased by the presence of mass transfers between phases that alter the equilibria in each phase.

The model-identification task can be carried out in either a simultaneous or an incremental manner. The simultaneous approach is based on a single identification step—all rate models are identified simultaneously, while the incremental approach decomposes the identification task into a set of subproblems of lower complexity—the rate laws and the model parameters are identified individually for each reaction.⁴ A detailed comparison between the simultaneous and incremental approaches has been proposed.⁵ Incremental model identification can be carried out using two different approaches, namely, via the rate-based method that relies on data differentiation⁶ or the extent-based method that uses data integration.⁴ In the rate-based method, the rates of the various dynamic processes are first computed by differentiation of measured concentrations, and each rate is subsequently modeled individually. In contrast, in the extent-based method, measured concentrations are first transformed to vessel extents, and each extent is then modeled individually by integration of the corresponding differential equation. These two incremental methods have been compared via a case study.⁷ The number of measured species required for transforming the

measured concentrations into rates or extents is the main limitation of incremental approaches. The possibility of using additional indirect measurements such as calorimetry⁸ and spectroscopy⁹ has also been documented.

Chemical reaction systems with instantaneous equilibria can be reduced to kinetically controlled reaction systems using quasi-steady state approximations via the method of singular perturbation.^{10,11} Once reduced, the reaction systems can be identified using either the simultaneous or the incremental approach. In the present work, instantaneous equilibria are treated in an exact way in the extent domain and combined with kinetically controlled reactions. Three different cases are considered for the extent-based incremental model identification of multiphase chemical reaction systems comprising instantaneous equilibria.

1. In the first case, the extents are computed directly from the measured concentrations of “all species” present in the reaction system. This yields extents for all reactions, that is, for both kinetically controlled and instantaneous equilibrium reactions, in addition to extents of mass transfers, inlets, and initial conditions. Note that the extents of equilibrium reactions capture the instantaneous shift in concentrations caused by the equilibria. This approach allows estimating the equilibrium constants from the measurements.

2. In the second case, the extents are computed from the concentrations of the “kinetic species” and the “equilibrium components”. These species represent the minimal set that is

Received: April 4, 2016

Revised: May 25, 2016

Accepted: May 26, 2016

Published: May 26, 2016

necessary to describe both the kinetically controlled and the equilibrium reactions. This leads to the computation of extents for the (slow) kinetically controlled reactions, but not for the instantaneous equilibria that are treated as algebraic equations. The number of measured species required in this formulation is reduced compared to the first case, but the knowledge of equilibrium constants is needed to reconstruct all species concentrations.

3. In the third case, the extents are computed from a mix of available measurements that includes some of the “kinetic species”, “equilibrium components”, and “equilibrium species”. In such a case, the reconstruction of all species concentrations requires the knowledge of equilibrium constants and the solution of nonlinear algebraic equations. Then, the transformation developed in the previous two cases can be applied to compute the extents.

This paper is organized as follows. Section 2 describes the material balance equations for the general case of two-phase reaction systems. After a description of various types of species and components, the balance equations are written first for all species (case 1 mentioned above), and then in terms of “kinetic species” and “equilibrium components” (case 2). Section 3 presents the corresponding transformations to vessel extents, including the case of a mix of available measurements (case 3). The extent-based incremental identification method for two-phase reaction systems in the presence of instantaneous equilibria is proposed in section 4 and illustrated in section 5 via a simulated example. Finally, section 6 concludes the paper.

2. MODEL OF TWO-PHASE CHEMICAL REACTION SYSTEMS

Consider a two-phase chemical reaction system with the phases G and L (for example, a gas and a liquid phase, or two liquid phases). The two phases are connected by p_m mass transfers. In addition, the reaction system has p_g inlets and one outlet in phase G, and p_l inlets and one outlet in phase L. The following assumptions are made:

1. The phases G and L are homogeneous.
2. The kinetically controlled reactions and the instantaneous equilibria occur only in the bulk of each phase.
3. The driving forces for mass transfer are the concentration gradients of the transferring species. Mass transfer is described by the steady-state film theory, and the flux is considered positive for a species transferring from phase G to phase L.

2.1. Species and Components. *Species.* The system has S species living in the S -dimensional set \mathcal{S} . Among these S species, S_f species are present in phase F and constitute the set \mathcal{S}_f , with $F \in \{G, L\}$, $f \in \{g, l\}$, and $S = S_g + S_l$. These S_f species are involved in R_f reactions in phase F, of which $R_{f,k}$ are kinetically controlled reactions and $R_{f,e}$ are instantaneous equilibria, with $R_f = R_{f,k} + R_{f,e}$. The subscripts $(\cdot)_k$ and $(\cdot)_e$ refer to kinetically controlled reactions and instantaneous equilibria, respectively. Note that if a species is present in both phases, it contributes to both S_g and S_l .

Kinetic and Equilibrium Species. The set \mathcal{S}_f can be partitioned into the set $\mathcal{S}_{f,k}$ of kinetic species and the complementary set $\mathcal{S}_{f,e}$ of equilibrium species. The $S_{f,k}$ kinetic species are only involved in the $R_{f,k}$ kinetically controlled reactions, that is, not in the instantaneous equilibria. The $S_{f,e}$ equilibrium species are involved in the $R_{f,e}$ instantaneous equilibria (and possibly in some of the kinetically controlled reactions). Hence, $\mathcal{S}_f = \mathcal{S}_{f,k} \cup \mathcal{S}_{f,e}$ and $S_f = S_{f,k} + S_{f,e}$.

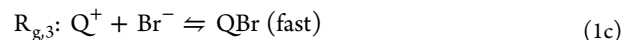
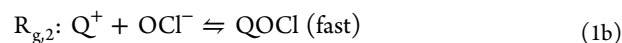
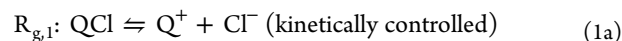
Equilibrium Components. An equilibrium component is a molecule constituent that is conserved by instantaneous equilibria. The set $\mathcal{S}_{f,c}$ contains the $S_{f,c}$ equilibrium components that are found in $\mathcal{S}_{f,e}$ and whose total numbers of moles are conserved in the instantaneous equilibria. The subscript $(\cdot)_c$ will be used to indicate an equilibrium component that is conserved by the instantaneous equilibria.

Mix of Available Measurements. To complete the terminology, let us define the set $\mathcal{S}_{f,a}$ describing a mix of $S_{f,a}$ available measurements that includes some of the kinetic species, equilibrium species, and equilibrium components.

Example. To illustrate the notation, let us consider the instantaneous equilibrium reaction $Q^+ + OCl^- \rightleftharpoons QOCl$ that involves three equilibrium species and the two equilibrium components labeled Q_c (including Q^+ and $QOCl$) and OCl_c (including OCl^- and $QOCl$). The reduction of the $S_{f,e}$ equilibrium species to the $S_{f,c}$ equilibrium components is a linear operation that can be represented by the matrix E_f of dimension $S_{f,c} \times S_{f,e}$.¹² The matrix E_f consisting of 0 and 1 elements indicates in which equilibrium species the various equilibrium component are present: a 1 in a given row and column indicates that the equilibrium component in that row is present in the equilibrium species of the corresponding column. In this simple example, with $\mathcal{S}_{f,e} = \{Q^+, OCl^-, QOCl\}$ and $\mathcal{S}_{f,c} = \{Q_c, OCl_c\}$,

$$E_f = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

To continue with this example, consider the aqueous phase (denoted as phase G) of the heterogeneous oxidation of benzyl alcohol with hypochlorite that will be described in the case study of section 5:



This reaction system in phase G consists of one kinetically controlled reaction (eq 1a) and two instantaneous equilibria (eqs 1b and 1c), thus giving $R_g = 3$, $R_{g,k} = 1$, and $R_{g,e} = 2$. The set of species in this phase is $\mathcal{S}_g = \{Cl^-, QCl, Q^+, OCl^-, Br^-, QOCl, OBr\}$, with the set of kinetic species $\mathcal{S}_{g,k} = \{Cl^-, QCl\}$ and the set of equilibrium species $\mathcal{S}_{g,e} = \{Q^+, OCl^-, Br^-, QOCl, QBr\}$. The set of components $\mathcal{S}_{g,c} = \{Q_c, OCl_c, Br_c\}$ can be constructed from $\mathcal{S}_{g,e}$ using the (3×5) matrix E_g

$$E_g = \begin{matrix} & Q^+ & OCl^- & Br^- & QOCl & QBr \\ \begin{matrix} Q_c \\ OCl_c \\ Br_c \end{matrix} & \begin{bmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix} \end{matrix}$$

where the element $E_g(i,j)$ is the number of times the equilibrium component i appears in the equilibrium species j .

For a reaction system consisting of both kinetically controlled and equilibrium reactions, the balance equations for phase F can be written in two ways, namely, in terms of the set of all species

S_f or in terms of the kinetic species $S_{f,k}$ and the equilibrium components $S_{f,c}$, as discussed in the next two subsections.

2.2. Balance Equations for All Species. The mole balance equations for all species S_f in phase F are¹³

$$\dot{\mathbf{n}}_f(t) = \mathbf{N}_f^T \mathbf{r}_{v,f}(t) \pm \mathbf{W}_{m,f} \boldsymbol{\zeta}(t) + \mathbf{W}_{in,f} \mathbf{u}_{in,f}(t) - \omega_f(t) \mathbf{n}_f(t), \quad \mathbf{n}_f(0) = \mathbf{n}_{f0} \quad (2)$$

where \mathbf{n}_f is the S_f -dimensional vector of numbers of moles, \mathbf{N}_f^T is the $S_f \times R_f$ stoichiometric matrix, $\mathbf{r}_{v,f}(t) := V_f(t) \mathbf{r}_f(t)$ is the R_f -dimensional vector of reaction rates, with V_f the volume of phase F, $\mathbf{W}_{in,f}$ the $S_f \times p_f$ inlet matrix expressing the composition of the inlets to phase F, $\mathbf{u}_{in,f}$ the p_f -dimensional vector of inlet mass flow rates to phase F, $\omega_f(t) := u_{out,f}(t)/m_f(t)$ is the inverse residence time, with $u_{out,f}$ the outlet mass flow rate and m_f the mass of phase F, and \mathbf{n}_{f0} is the vector of initial numbers of moles. The p_m mass transfers are treated as pseudo-inlets with unknown flow rates and involve the $S_f \times p_m$ mass-transfer matrix $\mathbf{W}_{m,p}$ where in each column the elements corresponding to a transferring species are one and the other elements are zero, and $\boldsymbol{\zeta}$ is the p_m -dimensional vector of mass-transfer rates expressed in moles/time units. By convention, a positive sign is assigned to mass transfers from phase G to phase L, and a negative sign in the other direction, which leads to the signs $-\mathbf{W}_{m,g} \boldsymbol{\zeta}(t)$ for phase G and $+\mathbf{W}_{m,l} \boldsymbol{\zeta}(t)$ for phase L. Note that the reaction rate vector $\mathbf{r}_{v,f}$ is of dimension R_f that is, it includes both the kinetically controlled and the equilibrium reactions. Concentrations can be computed as $\mathbf{c}_f(t) := \mathbf{n}_f(t)/V_f(t)$.

2.3. Balance Equations for Kinetic Species and Equilibrium Components. From $S_f = S_{f,k} + S_{f,e}$, the $(S_f \times R_f)$ -dimensional stoichiometric matrix \mathbf{N}_f^T can be partitioned into two submatrices, the $(S_{f,k} \times R_f)$ -dimensional matrix $\mathbf{N}_{f,k}^T$ associated with the $S_{f,k}$ kinetic species and the $(S_{f,e} \times R_f)$ -dimensional matrix $\mathbf{N}_{f,e}^T$ associated with the $S_{f,e}$ equilibrium species:

$$\mathbf{N}_f^T = \begin{bmatrix} \mathbf{N}_{f,k}^T \\ \mathbf{N}_{f,e}^T \end{bmatrix} \quad (3)$$

Similarly, the matrices $\mathbf{W}_{in,f}$ and $\mathbf{W}_{m,f}$ are partitioned into submatrices that are associated with the $S_{f,k}$ kinetic species, $\mathbf{W}_{in,f,k}$ and $\mathbf{W}_{m,f,k}$, and the $S_{f,e}$ equilibrium components, $\mathbf{W}_{in,f,e}$ and $\mathbf{W}_{m,f,e}$:

$$\mathbf{W}_{in,f} = \begin{bmatrix} \mathbf{W}_{in,f,k} \\ \mathbf{W}_{in,f,e} \end{bmatrix}, \quad \mathbf{W}_{m,f} = \begin{bmatrix} \mathbf{W}_{m,f,k} \\ \mathbf{W}_{m,f,e} \end{bmatrix} \quad (4)$$

Considering that the numbers of moles of the equilibrium components can be written as $\mathbf{n}_{f,c}(t) = \mathbf{E}_f \mathbf{n}_{f,e}(t)$, the mole balance equations for the $S_{f,k}$ kinetic species and the $S_{f,c}$ components read:

$$\underbrace{\begin{bmatrix} \dot{\mathbf{n}}_{f,k}(t) \\ \dot{\mathbf{n}}_{f,c}(t) \end{bmatrix}}_{\dot{\mathbf{n}}_f(t)} = \underbrace{\begin{bmatrix} \mathbf{N}_{f,k}^T \\ \mathbf{E}_f \mathbf{N}_{f,e}^T \end{bmatrix}}_{\mathbf{N}_f^T} \mathbf{r}_{v,f}(t) \pm \underbrace{\begin{bmatrix} \mathbf{W}_{m,f,k} \\ \mathbf{E}_f \mathbf{W}_{m,f,e} \end{bmatrix}}_{\mathbf{W}_{m,f}} \boldsymbol{\zeta}(t) + \underbrace{\begin{bmatrix} \mathbf{W}_{in,f,k} \\ \mathbf{E}_f \mathbf{W}_{in,f,e} \end{bmatrix}}_{\mathbf{W}_{in,f}} \mathbf{u}_{in,f}(t) - \omega_f(t) \underbrace{\begin{bmatrix} \mathbf{n}_{f,k}(t) \\ \mathbf{n}_{f,c}(t) \end{bmatrix}}_{\mathbf{n}_f(t)} \quad (5)$$

where $\mathbf{n}_{f,k}$ is the $S_{f,k}$ -dimensional vector of numbers of moles of the kinetic species, $\mathbf{n}_{f,c}$ is the $S_{f,c}$ -dimensional vector of numbers of moles of the equilibrium components, and (\cdot) indicates a quantity of dimension $\bar{S}_f := S_{f,k} + S_{f,c} \leq S_f$. The initial conditions for eq 5 are

$$\bar{\mathbf{n}}_f(0) = \begin{bmatrix} \mathbf{n}_{f,k}(0) \\ \mathbf{n}_{f,c}(0) \end{bmatrix} = \begin{bmatrix} \mathbf{n}_{f,k0} \\ \mathbf{n}_{f,c0} \end{bmatrix}$$

The matrix $\bar{\mathbf{N}}_f^T$ of dimension $\bar{S}_f \times R_f$ has a special structure with only zeros in the $R_{f,e}$ columns corresponding to the instantaneous equilibria. This follows from the way the kinetic species and the equilibrium components have been chosen, namely, (i) the $S_{f,k}$ kinetic species are only involved in the $R_{f,k}$ kinetically controlled reactions and not in the instantaneous equilibria and (ii) the $S_{f,c}$ equilibrium components are conserved by the instantaneous equilibria, and therefore all their corresponding stoichiometric coefficients are zero. It follows that $\bar{\mathbf{N}}_f^T$ has rank $R_{f,k} \leq R_f$ and the columns corresponding to the instantaneous equilibria can be discarded. The resulting matrix $\bar{\mathbf{N}}_{f,k}^T$ has dimension $\bar{S}_f \times R_{f,k}$ and the vector $\mathbf{r}_{v,f}$ reduces to $\mathbf{r}_{v,f,k}$ of dimension $R_{f,k}$. The mole balance eq 5 becomes

$$\dot{\bar{\mathbf{n}}}_f(t) = \bar{\mathbf{N}}_{f,k}^T \mathbf{r}_{v,f,k}(t) \pm \bar{\mathbf{W}}_{m,f} \boldsymbol{\zeta}(t) + \bar{\mathbf{W}}_{in,f} \mathbf{u}_{in,f}(t) - \omega_f(t) \bar{\mathbf{n}}_f(t), \quad \bar{\mathbf{n}}_f(0) = \bar{\mathbf{n}}_{f0} \quad (6)$$

Note that because $R_{f,e}$ columns of $\bar{\mathbf{N}}_f^T$ are $\mathbf{0}_{\bar{S}_f}$, it follows from the rank-nullity theorem that $S_{f,c} + R_{f,e} = S_{f,e}$.

To illustrate the matrix notation, consider again the reaction system given by eqs 1a–1c, with $R_g = 3$, $R_{g,k} = 1$, $R_{g,e} = 2$, and $S_g^T = 7$, $S_{g,k} = 2$, $S_{g,e} = 5$, and $S_{g,c} = 3$. The stoichiometric matrix \mathbf{N}_g^T with respect to all species (Cl^- , QCl , Q^+ , OCl^- , Br^- , QOCl , QBr) can be written as

$$\mathbf{N}_g^T = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \\ 1 & -1 & -1 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The corresponding stoichiometric matrix in terms of the kinetic species and the equilibrium components is

$$\bar{\mathbf{N}}_g^T = \begin{bmatrix} \mathbf{N}_{g,k}^T \\ \mathbf{E}_g \mathbf{N}_{g,e}^T \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

The last two columns of the matrix $\bar{\mathbf{N}}_g^T$ corresponding to the equilibrium reactions are all zeros and hence can be discarded, thus giving:

$$\bar{\mathbf{N}}_{g,k}^T = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

3. TRANSFORMATION TO VESSEL EXTENTS

Extent-based incremental identification encompasses two steps; namely (i) the measured numbers of moles are transformed to vessel extents, and (ii) each extent is used individually to identify the corresponding rate law and parameters.⁴ This section recalls the linear transformation to vessel extents, first in terms of the numbers of moles of all species and then in terms of the numbers of moles of the kinetic species and the equilibrium components. The choice of the transformation depends on the number and the nature of the available measurements. Then, the general case of a mix of available measurements is considered. Finally, a transformation that uses flow rate information to reduce the number of required measured species is also presented.

3.1. Using Measurements of All Species. Let $\text{rank}([\mathbf{N}_f^T \pm \mathbf{W}_{m,f} \mathbf{W}_{in,f} \mathbf{n}_{f0}]) = d_f$ with $d_f := R_f + p_m + p_f + 1$, and the $(S_f \times q_f)$ -dimensional matrix \mathbf{P}_f describe the null space of the matrix $[\mathbf{N}_f^T \pm \mathbf{W}_{m,f} \mathbf{W}_{in,f} \mathbf{n}_{f0}]^T$, with $q_f := S_f - d_f$. Then, the transformation $\mathcal{T}_f = [\mathbf{N}_f^T \pm \mathbf{W}_{m,f} \mathbf{W}_{in,f} \mathbf{n}_{f0} \mathbf{P}_f]^{-1}$ converts the numbers of moles \mathbf{n}_f into R_f extents of reaction $\mathbf{x}_{r,f}$, p_m extents of mass transfer $\mathbf{x}_{m,f}$, p_f extents of inlet $\mathbf{x}_{in,f}$, one extent of initial conditions $x_{ic,f}$, and q_f invariants $\mathbf{x}_{iv,f}$ that are identically equal to zero:

$$\begin{bmatrix} \mathbf{x}_{r,f}(t) \\ \mathbf{x}_{m,f}(t) \\ \mathbf{x}_{in,f}(t) \\ x_{ic,f}(t) \\ \mathbf{x}_{iv,f}(t) \end{bmatrix} = \mathcal{T}_f \mathbf{n}_f(t) \quad (7)$$

The transformation \mathcal{T}_f transforms eq 2 to the following system of variant and invariant states:

$$\begin{aligned} \dot{\mathbf{x}}_{r,f}(t) &= \mathbf{r}_{v,f}(t) - \omega_f(t) \mathbf{x}_{r,f}(t), & \mathbf{x}_{r,f}(0) &= \mathbf{0}_{R_f} \\ \dot{\mathbf{x}}_{m,f}(t) &= \boldsymbol{\zeta}(t) - \omega_f(t) \mathbf{x}_{m,f}(t), & \mathbf{x}_{m,f}(0) &= \mathbf{0}_{p_m} \\ \dot{\mathbf{x}}_{in,f}(t) &= \mathbf{u}_{in,f}(t) - \omega_f(t) \mathbf{x}_{in,f}(t), & \mathbf{x}_{in,f}(0) &= \mathbf{0}_{p_f} \\ \dot{x}_{ic,f}(t) &= -\omega_f(t) x_{ic,f}(t), & x_{ic,f}(0) &= 1 \\ \mathbf{x}_{iv,f}(t) &= \mathbf{0}_{q_f} \end{aligned} \quad (8)$$

The dynamic expressions in eq 8 represent the *vessel extents* associated with the reactions $\mathbf{r}_{v,f}(t)$, the mass transfers $\boldsymbol{\zeta}(t)$, the inlet flows $\mathbf{u}_{in,f}(t)$, and the initial conditions \mathbf{n}_{f0} . The vessel extents represent the amount of material produced, consumed, or handled by a particular rate process that is still in the reactor at time t , the negative terms on the right-hand side of eq 8 expressing the amount of material that has left the reactor via the outlet stream. The invariants express relationships between numbers of moles that must be met at all times. A detailed description and interpretation of the transformation is available in Rodrigues et al.¹⁴

Using $\mathcal{T}_f^{-1} = [\mathbf{N}_f^T \pm \mathbf{W}_{m,f} \mathbf{W}_{in,f} \mathbf{n}_{f0} \mathbf{P}_f]$, and taking into account the fact that the invariants are identically equal to zero, the numbers of moles can be reconstructed as

$$\mathbf{n}_f(t) = \mathbf{N}_f^T \mathbf{x}_{r,f}(t) \pm \mathbf{W}_{m,f} \mathbf{x}_{m,f}(t) + \mathbf{W}_{in,f} \mathbf{x}_{in,f}(t) + \mathbf{n}_{f0} x_{ic,f}(t) \quad (9)$$

Remark 1. The $R_{f,k}$ extents of reaction $\mathbf{x}_{r,f,k}$ capture the effect of the kinetically controlled reactions, whereas the remaining $R_{f,e}$ extents $\mathbf{x}_{r,f,e}$ describe the effect of the instantaneous thermodynamic equilibria on the numbers of moles of each species, thus

$$\mathbf{x}_{r,f} = \begin{bmatrix} \mathbf{x}_{r,f,k} \\ \mathbf{x}_{r,f,e} \end{bmatrix}$$

Remark 2. The mole balance eq 2 contains q_f invariants, which can be expressed as algebraic relationships:

$$\mathbf{P}_f^T \mathbf{n}_f(t) = \mathbf{0}_{q_f} \quad (10)$$

For batch and semi-batch reactors, the numbers of invariants are $q_f = S_f - R_f - p_m$ and $q_f = S_f - R_f - p_m - p_b$ respectively, with the invariant relationships being $\mathbf{P}_f^T (\mathbf{n}_f(t) - \mathbf{n}_{f0}) = \mathbf{0}_{q_f}$ (see ref 15 for details).

These invariant relationships can be used for data reconciliation.^{15,16}

3.2. Using Measurements of Kinetic Species and Equilibrium Components. Let $\text{rank}([\bar{\mathbf{N}}_{f,k}^T \pm \bar{\mathbf{W}}_{m,f} \bar{\mathbf{W}}_{in,f} \bar{\mathbf{n}}_{f0}]) = \bar{d}_f$ with $\bar{d}_f := R_{f,k} + p_m + p_f + 1$, and the $(\bar{S}_f \times \bar{q}_f)$ matrix $\bar{\mathbf{P}}_f$ denote the null space of $[\bar{\mathbf{N}}_{f,k}^T \pm \bar{\mathbf{W}}_{m,f} \bar{\mathbf{W}}_{in,f} \bar{\mathbf{n}}_{f0}]^T$, with $\bar{q}_f := \bar{S}_f - \bar{d}_f$. Note that $\bar{q}_f = q_f$ since $S_f = S_{f,k} + S_{f,e}$, $R_f = R_{f,k} + R_{f,e}$, and $S_{f,c} + R_{f,e} = S_{f,e}$. Then, the matrix $\bar{\mathcal{T}}_f = [\bar{\mathbf{N}}_{f,k}^T \pm \bar{\mathbf{W}}_{m,f} \bar{\mathbf{W}}_{in,f} \bar{\mathbf{n}}_{f0} \bar{\mathbf{P}}_f]^{-1}$ transforms the \bar{S}_f -dimensional vector of numbers of moles $\bar{\mathbf{n}}_f$ of the kinetic species and equilibrium components into $R_{f,k}$ extents of kinetically controlled reaction $\mathbf{x}_{r,f,k}$, p_m extents of mass transfer $\mathbf{x}_{m,f}$, p_f extents of inlet $\mathbf{x}_{in,f}$, one extent of initial conditions $x_{ic,f}$, and \bar{q}_f invariants $\bar{\mathbf{x}}_{iv,f}$ that are identically equal to zero, as

$$\begin{bmatrix} \mathbf{x}_{r,f,k}(t) \\ \mathbf{x}_{m,f}(t) \\ \mathbf{x}_{in,f}(t) \\ x_{ic,f}(t) \\ \bar{\mathbf{x}}_{iv,f}(t) \end{bmatrix} = \bar{\mathcal{T}}_f \bar{\mathbf{n}}_f(t) \quad (11)$$

The transformed system reads

$$\begin{aligned} \dot{\mathbf{x}}_{r,f,k}(t) &= \mathbf{r}_{v,f,k}(t) - \omega_f(t) \mathbf{x}_{r,f,k}(t), & \mathbf{x}_{r,f,k}(0) &= \mathbf{0}_{R_{f,k}} \\ \dot{\mathbf{x}}_{m,f}(t) &= \boldsymbol{\zeta}(t) - \omega_f(t) \mathbf{x}_{m,f}(t), & \mathbf{x}_{m,f}(0) &= \mathbf{0}_{p_m} \\ \dot{\mathbf{x}}_{in,f}(t) &= \mathbf{u}_{in,f}(t) - \omega_f(t) \mathbf{x}_{in,f}(t), & \mathbf{x}_{in,f}(0) &= \mathbf{0}_{p_f} \\ \dot{x}_{ic,f}(t) &= -\omega_f(t) x_{ic,f}(t), & x_{ic,f}(0) &= 1 \\ \bar{\mathbf{x}}_{iv,f}(t) &= \mathbf{0}_{\bar{q}_f} \end{aligned} \quad (12)$$

Compared to the transformation matrix \mathcal{T}_f , the matrix $\bar{\mathcal{T}}_f$ computes $R_{f,k}$ reaction extents, as it only extracts extents corresponding to the kinetically controlled reactions. The numbers of moles of the kinetic species and equilibrium components can then be reconstructed as

$$\bar{\mathbf{n}}_f(t) = \bar{\mathbf{N}}_{f,k}^T \mathbf{x}_{r,f,k}(t) \pm \bar{\mathbf{W}}_{m,f} \mathbf{x}_{m,f}(t) + \bar{\mathbf{W}}_{in,f} \mathbf{x}_{in,f}(t) + \bar{\mathbf{n}}_{f0} x_{ic,f}(t) \quad (13)$$

Remark 3. The transformation of kinetic species and equilibrium components to extents does not require the knowledge of equilibrium constants. However, as will be

discussed in section 4.1, the reconstruction of all species from the measurements of the kinetic species and equilibrium components will require the knowledge of these constants.

Remark 4. The mole balance in eq 5 contains \bar{q}_f invariants, which can be expressed as the algebraic relationships.

$$\bar{\mathbf{P}}_f^T \bar{\mathbf{n}}_f(t) = \mathbf{0}_{\bar{q}_f} \quad (14)$$

For batch and semibatch reactors, the numbers of invariants are $\bar{q}_f = S_f - R_{f,k} - p_m$ and $\bar{q}_f = S_f - R_{f,k} - p_m - p_b$ respectively, with the invariant relationships being $\bar{\mathbf{P}}_f^T (\bar{\mathbf{n}}_f(t) - \bar{\mathbf{n}}_{f0}) = \mathbf{0}_{\bar{q}_f}$ (see ref 15 for details).

3.3. Using a Mix of Available Measurements. For the cases where the two previous approaches are not feasible because of a lack of rank condition for the transformation (either $q_f < 0$ or $\bar{q}_f < 0$), it is possible to include measurements of equilibrium species to reconstruct the concentrations of all species. For this, assume that the measurements of at least $S_{f,a} = \bar{d}_f$ kinetic species, equilibrium components, and equilibrium species are available. Provided that the values of all $R_{f,e}$ equilibrium constants are known, the numbers of moles of all S_f species and $S_{f,c}$ equilibrium components can be computed by solving a set of nonlinear algebraic equations.

The number of relationships necessary to compute the numbers of moles of all S_f species and $S_{f,c}$ equilibrium components from the $S_{f,a}$ measurements can be assessed as follows:

- $S_f + S_{f,c}$ quantities to compute
- $S_{f,a}$ available measurements
- $S_f + S_{f,c} - S_{f,a} = R_{f,e} + S_{f,c} + q_f$ independent relationships, since $S_f = \bar{d}_f + q_f = \bar{d}_f + R_{f,e} + q_f$ and $S_{f,a} = \bar{d}_f$.

These relationships can be put together as follows: First, we assume that $R_{f,e}$ equilibrium relationships are available. For the notation, let us define the value of the equilibrium constant of the j th instantaneous equilibrium in the phase F as $K_{f,j}$ and its model structure as $g_{f,j}(\cdot)$ and group all these values and structures into the vectors \mathbf{K}_f and $\mathbf{g}_f(\cdot)$, respectively. Second, there are $S_{f,c}$ numbers of moles of equilibrium components that can be computed from the knowledge of matrix \mathbf{E}_f which converts the equilibrium species into the equilibrium components. Third, there are q_f invariant relationships defined by eq 10 that can be used. Hence, there are as many equations as unknowns, and the following system of $R_{f,e} + S_{f,c} + q_f$ nonlinear equations can be solved for the $R_{f,e} + S_{f,c} + q_f$ unknown numbers of moles that constitute the set complementary to $S_{f,a}$.

$$\mathbf{g}_f \left(\frac{\mathbf{n}_{f,e}(t)}{V(t)} \right) - \mathbf{K}_f = \mathbf{0}_{R_{f,e}} \quad R_{f,e} \text{ equilibrium relations} \quad (15a)$$

$$\mathbf{n}_{f,c}(t) - \mathbf{E}_f \mathbf{n}_{f,e}(t) = \mathbf{0}_{S_{f,c}} \quad S_{f,c} \text{ equilibrium components} \quad (15b)$$

$$\bar{\mathbf{P}}_f^T \bar{\mathbf{n}}_f(t) = \mathbf{0} \quad q_f \text{ invariant relationships} \quad (15c)$$

If $S_{f,a} > \bar{d}_f$ the system of eqs 15 is solved in the least-squares sense, since there are more equations than unknowns. Note that the invariant relationships for batch and semibatch reactors read $\bar{\mathbf{P}}_f^T (\bar{\mathbf{n}}_f(t) - \bar{\mathbf{n}}_{f0}) = \mathbf{0}_{q_f}$.

This system of nonlinear algebraic equations can be solved for discrete time instants using the Newton–Raphson algorithm.^{12,17} Then, the extents can be computed either in terms of all species (using the transformation of section 3.1) or in terms of the kinetic species and equilibrium components (via the transformation of section 3.2). Note that, since $\bar{q}_f = q_b$ the

invariant relationships in terms of kinetic species and equilibrium components given by eq 14 can be used in place of eq 15.

4. EXTENT-BASED MODEL IDENTIFICATION

Once extents have been computed from measured numbers of moles (or concentrations) using any of the approaches described in section 3, model identification and parameter estimation can be performed for each reaction and each mass transfer individually. For the case where the numbers of moles of the kinetic species and equilibrium components are measured, the extents can be computed using the approach outlined in section 3.2. However, since the rate expressions remain functions of the equilibrium species, and not of the equilibrium components, this requires redistributing the numbers of moles of the components to the species using the equilibrium constants.

4.1. Computation of Concentrations of Equilibrium Species. To compute the concentrations of the equilibrium species from the concentrations of the equilibrium components,

the $S_{f,e}$ -dimensional vector of concentrations $\mathbf{c}_{f,e}(t) = \frac{\mathbf{n}_{f,e}(t)}{V_f(t)}$ is partitioned into *independent* and *dependent* equilibrium species, whose concentrations are $\mathbf{c}_{f,e,i}$ (of dimension $S_{f,e,i}$) and $\mathbf{c}_{f,e,d}$ (of dimension $S_{f,e,d}$), respectively, with the number of independent equilibrium species being equal to the number of equilibrium components, $S_{f,e,i} = S_{f,c}$. The concentrations of the dependent species are computed as functions of the concentrations of the independent species via the knowledge of the equilibrium constants.

The concentrations of the independent species $\mathbf{c}_{f,e,i}$ can be computed by forcing the component concentrations $\mathbf{E}_f \mathbf{c}_{f,e}(t)$ to match the measured values $\tilde{\mathbf{c}}_{f,c}(t)$, under $S_{f,e,d}$ nonlinear constraints $\mathbf{h}_f(\cdot)$ corresponding to the known equilibrium relationships. The expressions of the equilibrium constants \mathbf{q}_f are rearranged to explicit concentrations of the dependent equilibrium species $\mathbf{c}_{f,e,d}(t) = \mathbf{h}_f(\mathbf{c}_{f,e,i}(t), K_{f,1}, \dots, K_{f,R_{f,e}})$.

$$\tilde{\mathbf{c}}_{f,c}(t) - \mathbf{E}_f \begin{bmatrix} \mathbf{c}_{f,e,i}(t) \\ \mathbf{c}_{f,e,d}(t) \end{bmatrix} = \mathbf{0}_{S_{f,c}} \quad \forall t \in \{t_1, \dots, t_H\}$$

$$\mathbf{c}_{f,e,d}(t) = \mathbf{h}_f(\mathbf{c}_{f,e,i}(t), K_{f,1}, \dots, K_{f,R_{f,e}}) \quad (16)$$

where the symbol $\tilde{(\cdot)}$ is used to define a measured quantity, $K_{f,j}$, $j = 1, \dots, R_{f,e}$, denotes the value of the j th equilibrium constant in phase F, the matrix \mathbf{E}_f of dimension $(S_{f,c} \times S_{f,e})$ is used to convert the equilibrium species into equilibrium components, and H is the number of measurements points. In practice, eq 16 is typically solved for $\mathbf{c}_{f,e,i}(t)$ using the Newton–Raphson algorithm.^{12,17}

4.2. Estimation of Kinetic Parameters. For each rate process to be modeled, a set of candidate models is postulated, and a nonlinear regression problem is solved for each of these candidates. For example, the regression problem to be solved for the i th reaction based on all concentration measurements is

$$\min_{\theta_{f,i}} \sum_{h=1}^H (\tilde{x}_{r,f,i}(t_h) - x_{r,f,i}(t_h, \theta_{f,i}))^2, \quad \forall i = 1, \dots, R_{f,k}$$

$$\text{s.t. } \dot{x}_{r,f,i}(t, \theta_{f,i}) = r_{v,f,i}(\tilde{\mathbf{c}}_f(t), \theta_{f,i}) - \omega_f(t) x_{r,f,i}(t, \theta_{f,i}),$$

$$x_{r,f,i}(0) = 0 \quad (17)$$

where θ_{fi} is a vector of uncertain parameters for the reaction rate $r_{v,fi}$ and $\mathbf{c}_f(t) = \mathbf{n}_f(t)/V_f(t)$ is the S_f -dimensional concentration vector. A similar problem can be formulated when working with the reduced set of kinetic species and equilibrium components, or when a mass-transfer expression has to be identified from its corresponding extent of mass transfer. Note that the parameter estimation problem has been formulated in the least-squares sense.¹⁸

4.3. Estimation of Equilibrium Constants. If the numbers of moles of all species are measured, it is possible to estimate the equilibrium constant for each instantaneous equilibrium. The equilibrium constant K_{fj} of the j th instantaneous equilibrium in phase F has the model structure $g_{fj}(\cdot)$. This model structure provides a constant value—the equilibrium constant—once the involved concentrations have reached their thermodynamic equilibrium. This equilibrium is fulfilled at all times since the instantaneous reactions are constantly at *quasi steady state*. Hence, it is possible to compute the value of the equilibrium constant of the j th instantaneous equilibrium at any time as $K_{fj}(t) = g_{fj}(\mathbf{c}_{f,e}(t))$.

One can compute an estimate of the equilibrium constant at time t_h as $\tilde{K}_{fj}(t_h) = g_{fj}(\tilde{\mathbf{c}}_{f,e}(t_h))$. In the absence of measurement noise, the value computed at time t_h is the same as the value computed at time t_{h+1} . However, since this is no longer true in the presence of noise, one typically computes a reliable estimate by averaging over the H time instants t_1, \dots, t_H . Such an estimator is statistically justified in the presence of zero-mean white noise.

To identify the equilibrium expression from a set of candidate models, one can select the expression with the minimal variance computed for the time instants t_1, \dots, t_H . The mean value of this model expression is an approximation of the equilibrium constant. For example, for the j th instantaneous reaction, given the two candidate models $g_{fj}^{(1)}(\cdot)$ and $g_{fj}^{(2)}(\cdot)$ and assuming $\text{Var}[g_{fj}^{(1)}(\tilde{\mathbf{c}}_{f,e}(t_h))] < \text{Var}[g_{fj}^{(2)}(\tilde{\mathbf{c}}_{f,e}(t_h))]$, one identifies the first model expression $g_{fj}^{(1)}(\cdot)$ and estimates the equilibrium constant as $\tilde{K}_{fj} = \frac{1}{H} \sum_{h=1}^H g_{fj}^{(1)}(\tilde{\mathbf{c}}_{f,e}(t_h))$.

In more general terms, the estimation of the equilibrium constant K_{fj} consists in finding, among the set of possible models $\mathcal{M}_j = \{g_{fj}^{(1)}, \dots, g_{fj}^{(M_j)}\}$ of dimension M_j , the pair of equilibrium constant and model structure $\tilde{K}_{fj}^*(t_h) = g_{fj}^{(m^*)}(\tilde{\mathbf{c}}_{f,e}(t_h))$ of minimal variance computed over the time instants $t_h \in \{t_1, \dots, t_H\}$:

$$m^* = \arg \min_m \text{Var}[g_{fj}^{(m)}(\tilde{\mathbf{c}}_{f,e}(t_h))] \quad \forall m = 1, \dots, M_j, g_{fj}^{(m)} \in \mathcal{M}_j, t_h \in \{t_1, \dots, t_H\} \quad (18)$$

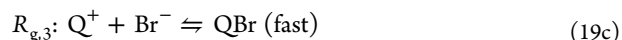
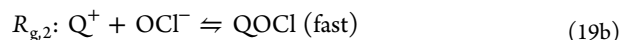
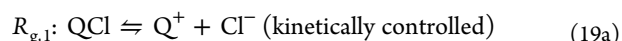
The equilibrium constant can be estimated from $g_{fj}^{(m^*)}(\cdot)$ as $\tilde{K}_{fj} = \frac{1}{H} \sum_{h=1}^H g_{fj}^{(m^*)}(\tilde{\mathbf{c}}_{f,e}(t_h))$.

5. CASE STUDY

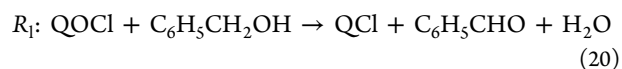
The application of the extent-based incremental identification to two-phase reaction systems is illustrated via the simulated example of the oxidation of benzyl alcohol by hypochlorite in a batch reactor.¹⁹ Three cases are considered, namely, (i) measurements of all species in both phases, (ii) measurements of the kinetic species and equilibrium components in the phase where the instantaneous equilibria take place (phase G), and (iii) measurements of a mix of kinetic species, equilibrium components, and equilibrium species in phase G.

5.1. Reaction System and Available Measurements.

The reaction system consists of an aqueous phase (labeled G) with one kinetically controlled reaction and two instantaneous equilibria ($R_g = 3, R_{g,k} = 1, R_{g,e} = 2$):



and an organic phase (labeled L) with a single kinetically controlled reaction ($R_l = R_{l,k} = 1$):



The volume of both phases is assumed to be constant and equal to 0.5 L. The mass-transfer resistance between the two phases is described by a thin film on each side of the interface, under quasi steady-state conditions. The characteristics of these two films—surface area, mass-transfer rate laws, and parameters—are assumed to be the same.

In the aqueous phase, cetyltrimethylammonium bromide (QBr) dissociates instantaneously to form the ions Q^+ and Br^- (eq 19c). The Q^+ ions react instantaneously with the hypochlorite ions (OCl^-) to form QOCl (eq 19b), which transfers to the organic phase. In the organic phase, benzyl alcohol ($\text{C}_6\text{H}_5\text{CH}_2\text{OH}$) in excess reacts with QOCl coming from the aqueous phase and forms benzyl aldehyde ($\text{C}_6\text{H}_5\text{CHO}$), QCl, and water. The species QCl formed in the organic phase transfers to the aqueous phase, where its dissociation in ions Q^+ and Cl^- (eq 19a) is kinetically observable.

The reaction and mass-transfer rate expressions as well as the equilibrium constants used for simulating the reaction schemes 19 and 20 are

$$r_g(t) = k_{g,1} c_{\text{QCl},g}(t) - k_{g,2} c_{\text{Q}^+,g}(t) c_{\text{Cl}^-,g}(t) \quad (21)$$

$$r_l(t) = k_1 c_{\text{QOCl},l}(t) \quad (22)$$

$$\zeta_{\text{QOCl}}(t) = k_{m,\text{QOCl}} A_1 V_1 (c_{\text{QOCl},g}(t) - c_{\text{QOCl},l}(t)) \quad (23)$$

$$\zeta_{\text{QCl}}(t) = k_{m,\text{QCl}} A_1 V_1 (c_{\text{QCl},g}(t) - c_{\text{QCl},l}(t)) \quad (24)$$

$$K_{g,2} = \frac{c_{\text{QOCl},g}(t)}{c_{\text{Q}^+,g}(t) c_{\text{OCl}^-,g}(t)} \quad (25)$$

$$K_{g,3} = \frac{c_{\text{QBr},g}(t)}{c_{\text{Q}^+,g}(t) c_{\text{Br}^-,g}(t)} \quad (26)$$

with A_1 being the specific interfacial area. The various sets of species present in this reaction system are listed in Table 1. Note that $S_{l,k} = S_l$ with $S_{l,e}$ and $S_{l,c}$ empty.

The stoichiometric matrices \mathbf{N}_g^T and \mathbf{N}_l^T and the mass-transfer matrices $\mathbf{W}_{m,g}$ and $\mathbf{W}_{m,l}$ are

Table 1. Sets of Species and Components Involved in the Oxidation of Benzyl Alcohol with Hypochlorite

set	species	dimension
S_1	{QOCl, C ₆ H ₅ CH ₂ OH, QCl, C ₆ H ₅ CHO, H ₂ O}	$S_1 = 5$
$S_{l,k}$	{QOCl, C ₆ H ₅ CH ₂ OH, QCl, C ₆ H ₅ CHO, H ₂ O}	$S_{l,k} = 5$
$S_{l,e}$	{∅}	$S_{l,e} = 0$
$S_{l,c}$	{∅}	$S_{l,c} = 0$
S_g	{Cl ⁻ , QCl, Q ⁺ , OCl ⁻ , Br ⁻ , QOCl, QBr }	$S_g = 7$
$S_{g,k}$	{Cl ⁻ , QCl }	$S_{g,k} = 2$
$S_{g,e}$	{Q ⁺ , OCl ⁻ , Br ⁻ , QOCl, QBr }	$S_{g,e} = 5$
$S_{g,c}$	{Q ⁺ , OCl ⁻ , Br ⁻ }	$S_{g,c} = 3$
S_m	{QOCl, QCl}	$S_m = 2$

$$N_g^T = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \\ 1 & -1 & -1 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad N_l^T = \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad W_{m,g} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}$$

$$W_{m,l} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

The $S_{g,e} = 5$ equilibrium species in the aqueous phase G can be reduced to $S_{g,c} = 3$ equilibrium components that represent the following conserved quantities:

$$c_{Q^+,g} := c_{Q^+,g} + c_{QOCl,g} + c_{QBr,g} \\ = c_{Q^+,g} (1 + K_{g,2}c_{OCl^-,g} + K_{g,3}c_{Br^-,g}) \quad (27a)$$

$$c_{OCl^-,g} := c_{OCl^-,g} + c_{QOCl,g} = c_{OCl^-,g} (1 + K_{g,2}c_{Q^+,g}) \quad (27b)$$

$$c_{Br^-,g} := c_{Br^-,g} + c_{QBr,g} = c_{Br^-,g} (1 + K_{g,3}c_{Q^+,g}) \quad (27c)$$

where the concentrations of the equilibrium products $c_{QOCl,g}$ and $c_{QBr,g}$ have been replaced by their expression obtained from the thermodynamic equilibrium (eqs 25 and 26). The values of the model parameters used in the simulation are given in Table 2.

The aqueous phase G is loaded with 0.125 kmol of OCl⁻ in ionic form, 0.04 kmol of QBr, and 0.005 kmol of QCl. The initial amount of benzyl alcohol in the organic phase L is 0.968 kmol. The numbers of moles in both phases are simulated for 10 min, and all the measured numbers of moles are corrupted with 2% zero-mean Gaussian noise with respect to the maximal number of moles of each species. Another set of data is obtained by corrupting the measured numbers of moles with 5% zero-mean

Gaussian noise to be used for parameter estimation in section 5.3.

5.2. Computation of Extents. For the computation of extents, three situations with the following measurements are considered: (case 1) all species in both phases, (case 2) all species in phase L and the kinetic and equilibrium species in phase G, and (case 3) all species in phase L and a mix of kinetic species, equilibrium components, and equilibrium species in phase G.

5.2.1. Case 1: Measurement of All Species. The extents are obtained from the measured numbers of moles using the transformations \mathcal{T}_1 and \mathcal{T}_g of section 3.1. For phase L, $R_l = 1$ extent of reaction and $p_m = 2$ extents of mass transfers are extracted from the $S_l = 5$ measured numbers of moles. For phase G, $R_{g,k} = 1$ extent of (kinetically controlled) reaction, $R_{g,e} = 2$ extents of equilibria, and $p_m = 2$ extents of mass transfers are extracted from the $S_g = 7$ measured numbers of moles. The measured numbers of moles as well as the extents of reaction and mass transfer are shown in Figure 1. Note that, since C₆H₅CH₂OH is in large excess, its concentration in the organic phase is constant and thus not shown. The amount of Q⁺ in the aqueous phase is also too small to be represented in Figure 1. In addition, since the reaction system has no outlet (batch conditions), the extents of mass transfer $x_{m,g}$ and $x_{m,l}$ are the same.

Minimal Number of Measured Species. The transformations of section 3.1 require the measurement of $R_l + p_m = 3$ species in the organic phase and $R_g + p_m = 5$ species in the aqueous phase. This condition can be relaxed by considering the dynamic relationship that exists between $x_{m,g}(t)$ and $x_{m,l}(t)$. In the case of steady-state mass transfer, $x_{m,g}(t)$ and $x_{m,l}(t)$ only differ due to $\omega_g(t)$ and $\omega_l(t)$. If this information is known or measured, the extents of mass transfer in one phase can be converted to the extents of mass transfer in the other phase.⁸ This allows reducing the number of required measured species by two, down to a minimum of six in total. Among these six species, $R_l = 1$ numbers of moles must be measured in the organic phase, while $R_{g,k} + R_{g,e} = 3$ numbers of moles must be measured in the aqueous phase, with the remaining $p_m = 2$ measured numbers of moles—required to compute the extents of mass transfer—coming from either of the two phases.

5.2.2. Case 2: Measurement of All Species in Phase L and of Kinetic Species and Equilibrium Components in Phase G. The extents are obtained using the transformation \mathcal{T}_1 of section 3.1 for the organic phase L and the transformation $\bar{\mathcal{T}}_g$ of section 3.2 for the aqueous phase G. For phase L, $R_l = 1$ extent of reaction and $p_m = 2$ extents of mass transfers are extracted from the $S_l = 5$ measured numbers of moles. For phase G, $R_{g,k} = 1$ extent of (kinetically controlled) reaction and $p_m = 2$ extents of mass transfers are computed from the $S_{g,k} + S_{g,c} = 5$ measured numbers of moles. Using the knowledge of equilibrium constants, the numbers of moles of the $S_{g,c} = 3$ components are redistributed into the $S_{g,e} = 5$ equilibrium species according to eq 16 to obtain the numbers of moles of the $S_g = 7$ species in

Table 2. Kinetic, Thermodynamic, and Surface Parameters Used in the Simulation

parameter	value	unit	parameter	value	unit
$k_{g,1}$	1.663×10^{-2}	s ⁻¹	k_i	22.7	s ⁻¹
$k_{g,2}$	2.5	m ³ kmol ⁻¹ s ⁻¹	$k_{m,QOCl}$	8.02×10^{-5}	m s ⁻¹
$K_{g,2}$	1.157×10^3	m ³ kmol ⁻¹	$k_{m,QCl}$	8.91×10^{-5}	m s ⁻¹
$K_{g,3}$	0.235×10^3	m ³ kmol ⁻¹	A_i	200	m ⁻¹

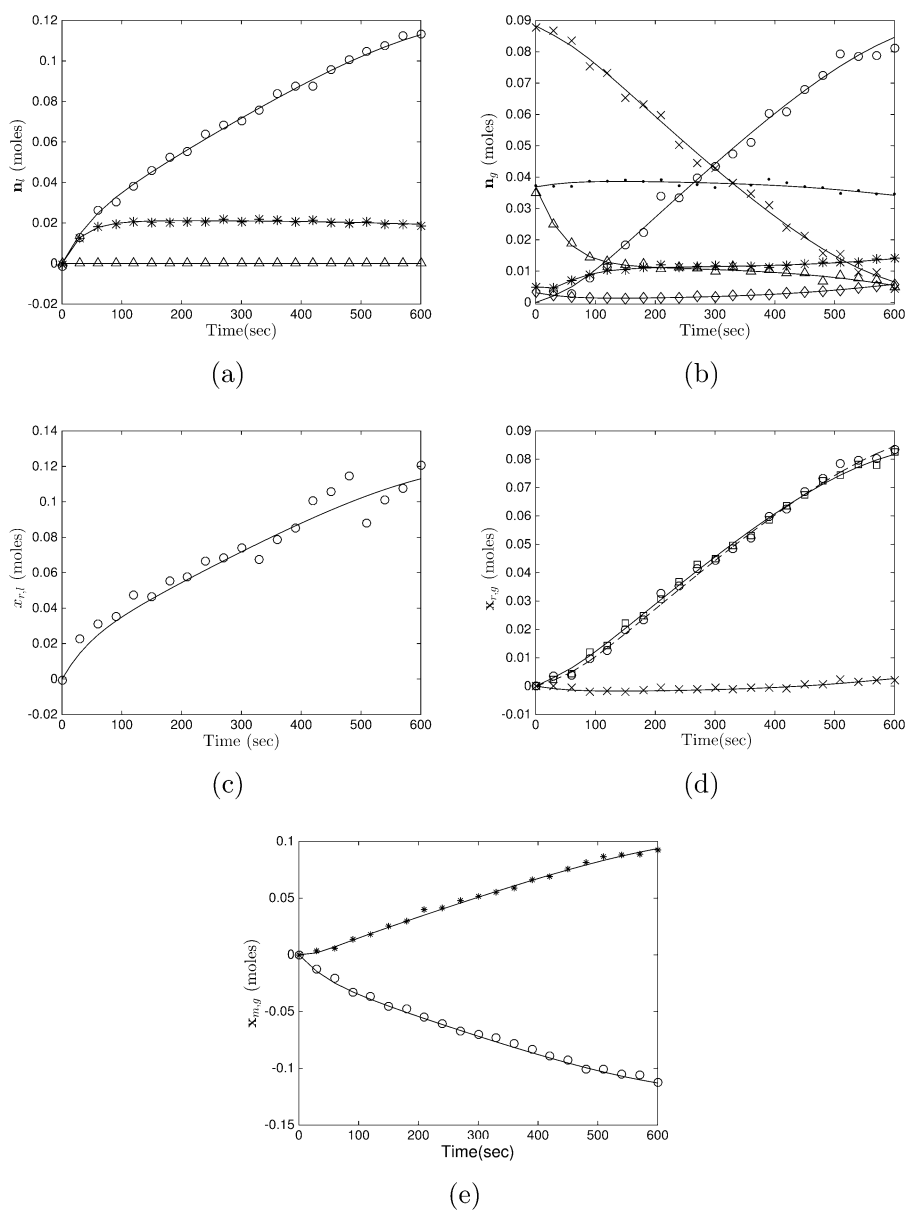


Figure 1. Case 1: Measurement of all species. (Top, a and b) simulated (noise-free, continuous lines) and measured (noisy, 2%) numbers of moles of species H_2O (O), QCl (*), and QOCl (Δ) in the organic phase and of species Cl^- (O), QCl (*), OCl^- (x), Br^- (•), QOCl (Δ), and QBr (\diamond) in the aqueous phase. (Center, c and d) Experimental (computed from measurements) and modeled (continuous and dashed lines) extents of reaction in the organic phase, with $x_{r,i}$ (O), and in the aqueous phase, with $x_{r,g,k}$ (x), $x_{r,g,e,1}$ (\square), and $x_{r,g,e,2}$ (O). (Bottom e) Experimental and modeled (continuous lines) extents of mass transfer of QOCl (*) and QCl (O) in the aqueous phase.

phase G. The measured numbers of moles and extents are shown in Figure 2.

Minimal Number of Measured Species. The use of the transformation of section 3.1 requires the measurement of $R_1 + p_m = 3$ species in the organic phase, whereas the transformation of section 3.2 requires $R_{g,k} + p_m = 3$ kinetic species and equilibrium components to be measured in the aqueous phase. As mentioned earlier, this requirement can be relaxed using the dynamic relationship between the extents of mass transfer in different phases, which reduces the number of required measured species by two down to a minimum of four in total. Among these four species, $R_1 = 1$ number of moles must be measured in the organic phase, while $R_{g,k} = 1$ number of moles must be measured in the aqueous phase, with the remaining $p_m = 2$ measured numbers of moles—required to compute the extents of mass transfer—coming from either of the two phases.

5.2.3. Case 3: Measurement of All Species in Phase L and of a Mix of Species in Phase G. Let the set of available measurements in the aqueous phase be $\{\text{Q}_{c,g}, \text{Q}_{g}^+, \text{QCl}_{g}\}$ of dimension $S_{g,a} = R_{g,k} + p_m$. Since this set consists neither of all S_g species nor of the $S_{g,k}$ kinetic species and the $S_{g,c}$ equilibrium components, the transformations \mathcal{T}_g and $\bar{\mathcal{T}}_g$ cannot be applied. Instead, the procedure described in section 3.3 is used to compute the numbers of moles of all $S_g = 7$ species and $S_{g,c} = 3$ equilibrium components, knowing the structure of the equilibrium relationships and the values of the equilibrium constants. The $R_{f,e} + S_{f,c} + q_f = 2 + 3 + 2 = 7$ equations used to compute the $S_g + S_{g,c} - S_{g,a} = 7 + 3 - 3 = 7$ unknown states at each time instant are as follows:

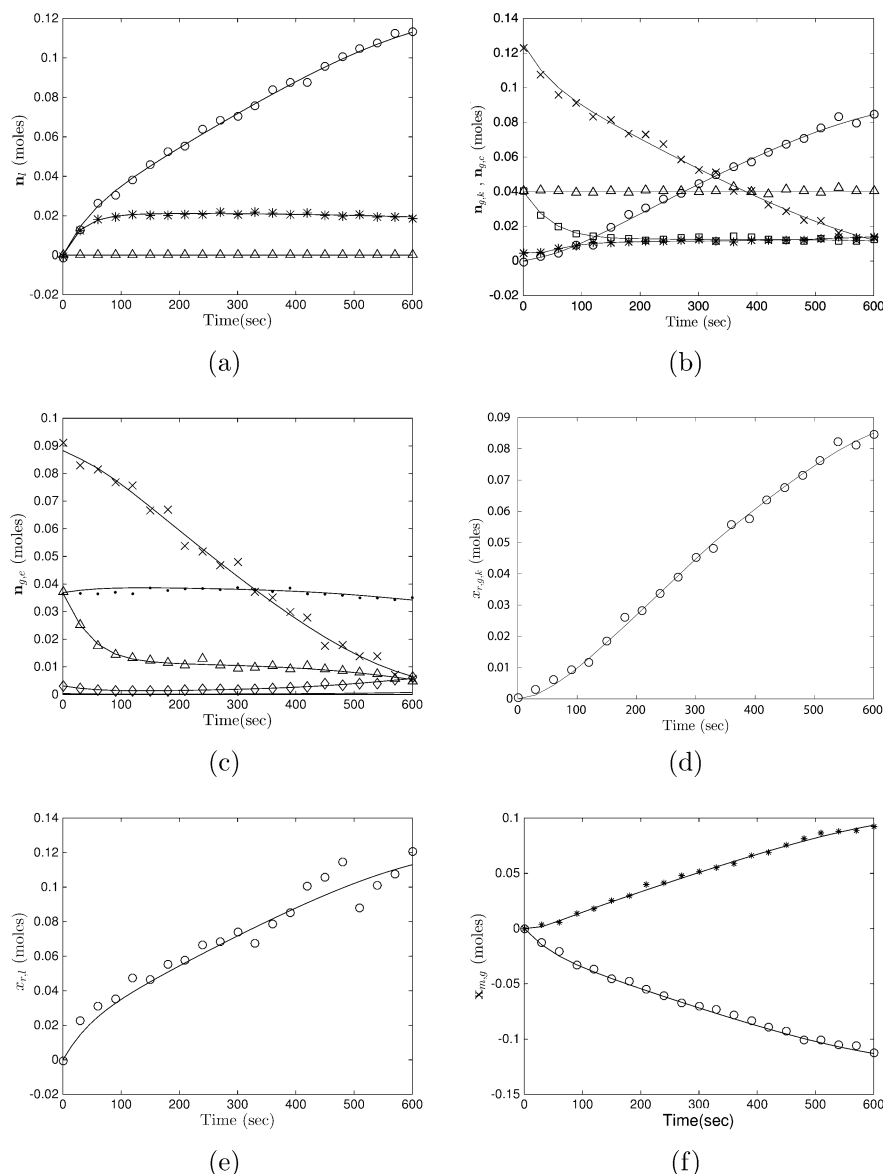


Figure 2. Case 2: Measurement of all species in phase L and of kinetic species and equilibrium components in phase G. (Top, a and b) Simulated (noise-free, continuous lines) and measured (noisy, 2%) numbers of moles of species H₂O (○), QCl (*) and QOCl (△) in the organic phase and of kinetic species Cl⁻ (○) and QCl (*) and equilibrium components Q_c (□), OCl_c (×), and Br_c (△) in the aqueous phase. (Center, c) Simulated (continuous lines) and reconstructed numbers of moles of the equilibrium species OCl⁻ (×), Br⁻ (●), QOCl (△), and QBr (◇) in the aqueous phase. (Center, d; bottom, e) Experimental (computed from measurements) and modeled (continuous lines) extents of (kinetically controlled) reaction in the aqueous and organic phases. (Bottom, f) Experimental and modeled (continuous lines) extents of mass transfer of species QOCl (*) and QCl (○) in the aqueous phase.

$$\frac{n_{QOCl,g}(t)}{n_{Q^+,g}(t) n_{OCl^-,g}(t)} V_g(t) - K_{g,2} = 0 \tag{28a}$$

$$\frac{n_{QBr,g}(t)}{n_{Q^+,g}(t) n_{Br^-,g}(t)} V_g(t) - K_{g,3} = 0 \tag{28b}$$

$$n_{Q^+,g}(t) - (n_{Q^+,g}(t) + n_{QOCl,g}(t) + n_{QBr,g}(t)) = 0 \tag{28c}$$

$$n_{OCl^-,g}(t) - (n_{OCl^-,g}(t) + n_{QOCl,g}(t)) = 0 \tag{28d}$$

$$n_{Br^-,g}(t) - (n_{Br^-,g}(t) + n_{QBr,g}(t)) = 0 \tag{28e}$$

$$n_{OCl^-,g}(t) + n_{Cl^-,g}(t) + n_{Br^-,g}(t) - n_{Q^+,g}(t) = n_{OCl^-,g0} + n_{Cl^-,g0} + n_{Br^-,g0} - n_{Q^+,g0} \tag{28f}$$

$$n_{Br^-,g}(t) + n_{QBr,g}(t) = n_{Br^-,g0} + n_{QBr,g0} \tag{28g}$$

Eqs 28a and 28b are equilibrium relations; eqs 28c, 28d, and 28e are derived from the definition of equilibrium components; and eqs 28f and 28g are invariant relationships under batch conditions (see comment on eq 10 in section 3.1). The measured and computed numbers of moles in the aqueous phase are shown in Figure 3. The extents can be computed using the transformation \mathcal{T}_g or $\bar{\mathcal{T}}_g$ (not shown).

5.3. Model Identification and Parameter Estimation. For the kinetically controlled reactions, the rate laws are

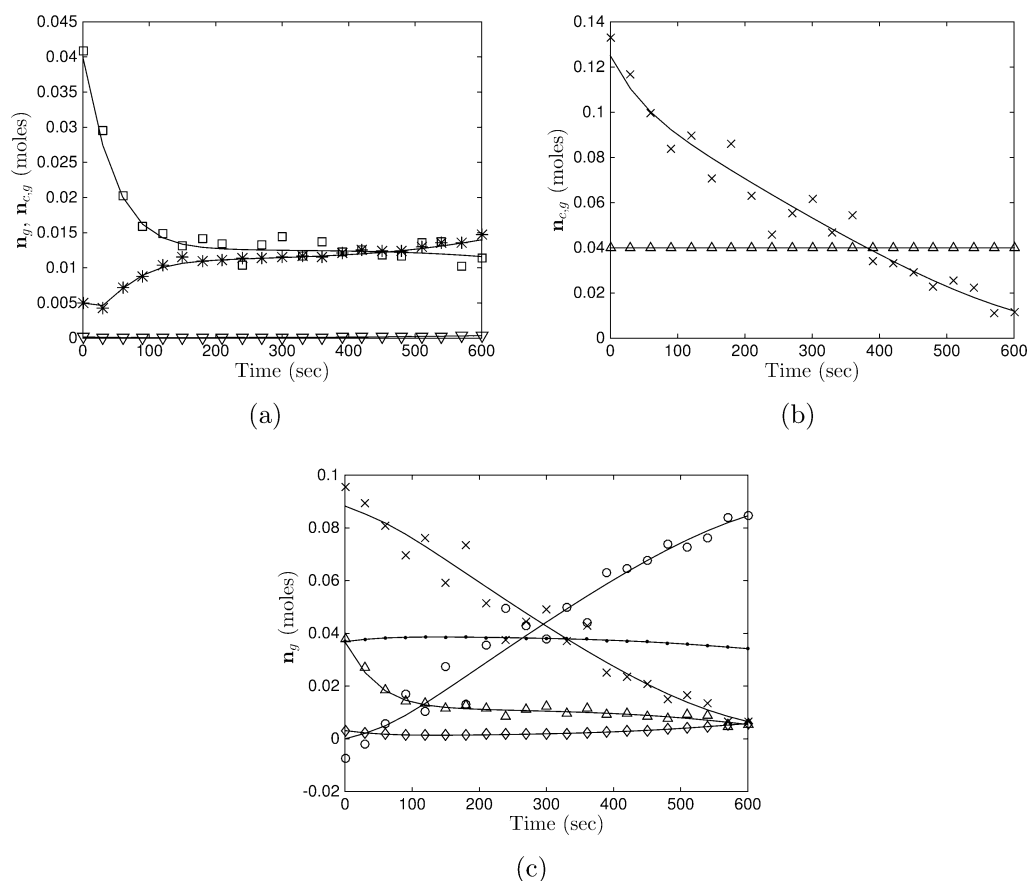


Figure 3. Case 3: Measurement of all species in phase L and of a mix of species in phase G. (Top, a and b) Simulated (noise-free, continuous lines) and measured (noisy, 2%) numbers of moles of component Q_c (\square), kinetic species QCl ($*$), and equilibrium species Q^+ (∇) in the aqueous phase (a) and reconstructed values of equilibrium components OCl_c (\times) and Br_c (\triangle) in the aqueous phase (b). (Bottom c) Simulated (continuous lines) and reconstructed numbers of moles of the species OCl^- (\times), Br^- (\bullet), $QOCl$ (\triangle), Cl^- (\circ), and QBr (\diamond) in the aqueous phase.

Table 3. Extent-Based Incremental Model Identification of the Kinetically Controlled Reactions in Phases G and L (2% noise), with the Corresponding Residual Sum of Squares (RSS)

$R_{g,1}$	rate expression	RSS		R_l	rate expression	RSS
		$x_{r,g}$ (case 1)	$x_{r,g,k}$ (case 2)			$x_{r,l}$
$r_g^{(1)}$	$k_{g,1}^{(1)} c_{Q^+} c_{Cl^-}$	2.01×10^{-2}	2.04×10^{-2}	$r_l^{(1)}$	$k_l^{(1)} c_{C_6H_5CH_2OH}$	3.4×10^{-3}
$r_g^{(2)}$	$k_{g,1}^{(2)} c_{QCl} - k_{g,2}^{(2)} c_{Q^+} c_{Cl^-}$	7.34×10^{-5}	1.12×10^{-4}	$r_l^{(2)}$	$k_l^{(2)} c_{QOCl}$	1.6×10^{-3}

Table 4. Estimated Rate Constants for Reactions $R_{g,1}$ and R_l and Mass-Transfer Coefficients for $QOCl$ and QCl , with their Corresponding 99% Confidence Intervals (C.I.) for Noise Levels of 2% and 5%

parameter	true value	fitted extent	2% noise [99% C.I.]	5% noise [99% C.I.]
$k_{g,1}$	1.663×10^{-2}	$x_{r,g}$	$1.62 [1.38-1.86] \times 10^{-2}$	$1.62 [1.08-2.16] \times 10^{-2}$
		$x_{r,g,k}$	$1.67 [1.58-1.76] \times 10^{-2}$	$1.74 [1.33-2.15] \times 10^{-2}$
$k_{g,2}$	2.5	$x_{r,g}$	$2.68 [1.13-4.25]$	$2.17 [0.00-5.20]$
		$x_{r,g,k}$	$2.64 [1.56-3.74]$	$2.71 [0.88-4.54]$
k_l	22.7	$x_{r,l}$	$23.70 [22.6-25.9]$	$24.30 [21.6-27.0]$
$k_{m,QOCl}$	8.02×10^{-5}	$x_{m,g}$	$7.90 [7.50-8.28] \times 10^{-5}$	$7.89 [6.54-9.24] \times 10^{-5}$
		$x_{m,l}$	$7.96 [7.43-8.49] \times 10^{-5}$	$8.10 [7.13-9.07] \times 10^{-5}$
$k_{m,QCl}$	8.91×10^{-5}	$x_{m,g}$	$8.83 [7.60-10.66] \times 10^{-5}$	$8.83 [8.42-9.27] \times 10^{-5}$
		$x_{m,l}$	$8.97 [8.46-9.48] \times 10^{-5}$	$9.04 [7.85-10.2] \times 10^{-5}$

identified from the extents $x_{r,l}$, $x_{r,g}$ (case 1), and $x_{r,g,k}$ (case 2) according to eq 17. The identification results are summarized in Table 3 for two competing models in each phase. Based on the residual sum of squares (RSS), the rate models $r_g^{(2)}$ and $r_l^{(2)}$, which correspond to the rate expressions 21 and 22 used in the simulation, are correctly identified. Note that the mass-transfer

rate expressions are assumed to be known, thus limiting the identification to the estimation of their parameters.

As shown in Table 4, the rate parameters $k_{g,1}$, $k_{g,2}$, k_l , $k_{m,QOCl}$, and $k_{m,QCl}$ are correctly estimated within 99% confidence regions. Figure 4 shows the resulting fit for the extent of reaction in the organic phase when the correct rate expression is used,

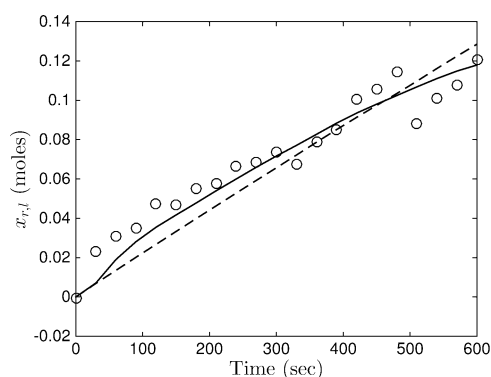


Figure 4. Experimental (2% noise, O) and modeled extent of reaction in the organic phase predicted using the rate expressions $r_1^{(2)}$ (—, correct model) and $r_1^{(1)}$ (---, incorrect model) of Table 3.

with the estimated value of k_1 for two different noise levels given in Table 4. For the aqueous phase, both fitting methods based on $x_{r,g}$ or $x_{r,g,k}$ lead to correct estimates of the two parameters. The small difference between these methods stems from the difference in the noise structure of $x_{r,g}$ and $x_{r,g,k}$, which results from the fact that all species are used to compute $x_{r,g}$ (case 1) but only the kinetic species and the equilibrium components to compute $x_{r,g,k}$ (case 2).

When a mix of species is measured (case 3), minor variations in the results compared to the other cases are observed (not shown). In addition, there are slight variations depending on the choice of the measured subset $S_{g,a}$. Note also that case 3 leads to two subcases depending on whether the transformation \mathcal{T}_g or $\bar{\mathcal{T}}_g$ is used to compute $x_{r,g}$ or $x_{r,g,k}$, respectively. There are also minor variations in the results due to differences in error propagation through the transformations \mathcal{T}_g and $\bar{\mathcal{T}}_g$ (not shown).

6. CONCLUSIONS

Modeling heterogeneous chemical reaction systems that comprise instantaneous equilibria has always been a difficult task. The presence of fast reactions—which can be modeled via instantaneous equilibria in the time scale of interest—affects the numbers of moles of other species, that is, also the other rate processes such as the kinetically controlled reactions and the mass transfers. Conventional modeling using the simultaneous approach requires modeling the effect of equilibria in addition to the other rate processes. In contrast, the incremental extent-based model identification technique simplifies the modeling by generating a single extent for each rate process. As a consequence, a kinetically controlled reaction can be modeled independently of the instantaneous equilibria.

This paper has investigated the use of incremental extent-based identification for two-phase reaction systems with instantaneous equilibria. Different alternatives for decoupling the reaction system based on the nature and the number of available measurements have been discussed. It has been shown that statistically correct estimation of the extents of reaction and mass transfer is possible from (1) measurements of all species, (2) measurements of *kinetic species* and *equilibrium components* only, and (3) measurement of a mix of *kinetic species*, *equilibrium components*, and *equilibrium species*. Case 1 allows estimating the equilibrium constants, whereas both cases 2 and 3 require the knowledge of equilibrium constants, case 2 for the identification

step and case 3 already for the transformation to extents. All these different cases have been illustrated via the simulated example of the oxidation of benzyl alcohol using hypochlorite in a batch reactor.

AUTHOR INFORMATION

Corresponding Author

*E-mail: dominique.bonvin@epfl.ch.

Notes

The authors declare no competing financial interest.

GLOSSARY

Scalars

- A = specific surface area
- d = dimensionality ($R + p + p_m + 1$)
- H = number of measured time instants
- K = equilibrium constant (evaluation of function g)
- k = rate constant
- M = number of equilibrium model expressions
- m = mass or index of equilibrium model (in section 4.3)
- p = number of independent inlet flows
- q = number of invariants ($S - d$)
- S = number of species
- R = number of independent reactions
- u_{out} = outlet mass flow rate
- V = volume of the reaction mixture
- ω = inverse residence time

Vectors Dimension

- c = concentrations ($S \times 1$)
- n = numbers of moles ($S \times 1$)
- r = reaction rates ($R \times 1$)
- u_{in} = inlet mass flow rates ($p \times 1$)
- x = vessel extents ($S \times 1$)
- θ = adjustable kinetic parameters ($n_\theta \times 1$)
- ζ = mass-transfer rates ($p_m \times 1$)

Matrices Dimension

- E = transformation of equilibrium species to components ($S_c \times S_e$)
- N = stoichiometry ($R \times S$)
- P = null space matrix ($S \times q$)
- W_{in} = inlet composition ($S \times p$)
- W_m = matrix of mass transfer ($S \times p_m$)
- \mathcal{T} = transformation of numbers of moles to extents ($S \times S$)

Subscripts

- a = available quantities
- c = equilibrium components
- d = dependent equilibrium species (as opposed to i)
- e = equilibrium species
- f = phase F, $F \in \{G, L\}$
- g = phase G
- h = specific time instant
- i = independent equilibrium species (as opposed to d)
- ic = initial conditions
- in = inlet flows
- iv = invariants
- k = kinetic species
- l = phase L
- m = mass transfers
- out = outlet flow
- r = kinetically controlled reactions
- v = premultiplication by the volume
- 0 = initial conditions

Superscripts

- T = vector or matrix transposition
 -1 = matrix inversion
 (i) = i th candidate model, with $i \in \mathbb{N}_+$
 $(\dot{\cdot})$ = derivative with respect to time
 $(\bar{\cdot})$ = reduction to kinetic species and equilibrium components
 $(\widetilde{\cdot})$ = measured quantity
 $*$ = value at the minimum

Sets

- \mathcal{M} = equilibrium model expressions, dimension: M
 \mathcal{S} = kinetic species, equilibrium species or components, dimension: S

Functions

- $g(\cdot)$ = equilibrium model expression
 $h(\cdot)$ = expression for the dependent equilibrium species (rearrangement of the function g , knowing K)
 $r(\cdot)$ = reaction rate expression
 $\text{Var}[\cdot]$ = variance function

REFERENCES

- Berger, R. J.; Stitt, E. H.; Marin, G.; Kapteijn, F.; Moulijn, J. A. Chemical Reaction Kinetics in Practice. *CATTECH* **2001**, *5*, 36.
- Norman, S.; Maeder, M. Model-Based Analysis for Kinetic and Equilibrium Investigations. *Crit. Rev. Anal. Chem.* **2006**, *36*, 199.
- Dyson, R.; Maeder, M.; Neuhold, Y. M.; Puxty, G. Analyses of Three-Way Data from Equilibrium and Kinetic Investigations. *Anal. Chim. Acta* **2003**, *490*, 99.
- 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.
- Bardow, A.; Marquardt, W. Incremental and Simultaneous Identification of Reaction Kinetics: Methods and Comparison. *Chem. Eng. Sci.* **2004**, *59*, 2673.
- Michalik, C.; Brendel, M.; Marquardt, W. Incremental Identification of Fluid Multi-Phase Reaction Systems. *AIChE J.* **2009**, *55*, 1009.
- Bhatt, N.; Kerimoglu, N.; Amrhein, M.; Marquardt, W.; Bonvin, D. Incremental Identification for Reaction Systems - A Comparison between Rate-Based and Extent-Based Approaches. *Chem. Eng. Sci.* **2012**, *83*, 24.
- Srinivasan, S.; Billeter, J.; Bonvin, D. Extent-Based Incremental Identification of Reaction Systems Using Concentration and Calorimetric Measurements. *Chem. Eng. J.* **2012**, *207–208*, 785.
- Billeter, J.; Srinivasan, S.; Bonvin, D. Extent-Based Kinetic Identification using Spectroscopic Measurements and Multivariate Calibration. *Anal. Chim. Acta* **2013**, *767*, 21.
- Vora, N.; Daoutidis, P. Nonlinear Model Reduction of Chemical Reaction Systems. *AIChE J.* **2001**, *47*, 2320.
- Kokotović, P.; Khalil, H. K.; O'Reilly, J. *Singular Perturbation Methods in Control: Analysis and Design*; Academic Press Inc.: London, 1986.
- Maeder, M.; Neuhold, Y. M. *Practical Data Analysis in Chemistry*; Elsevier: Amsterdam, 2006.
- Bhatt, N.; Amrhein, M.; Bonvin, D. Extents of Reaction, Mass Transfer and Flow for Gas–Liquid Reaction Systems. *Ind. Eng. Chem. Res.* **2010**, *49*, 7704.
- Rodrigues, D.; Srinivasan, S.; Billeter, J.; Bonvin, D. Variant and Invariant States for Reaction Systems. *Comput. Chem. Eng.* **2015**, *73*, 23.
- Srinivasan, S.; Billeter, J.; Narasimhan, S.; Bonvin, D. Data Reconciliation in Reaction Systems using the Concept of Extents. *Comput.-Aided Chem. Eng.* **2015**, *37*, 419.
- Narasimhan, S.; Jordache, C. *Data Reconciliation and Gross Error Detection*; Gulf Professional Publishing: Houston, TX, 1999.

(17) Maeder, M.; Neuhold, Y. M.; Puxty, G.; King, P. Analysis of Reactions in Aqueous Solution at Non-Constant pH: No More buffers? *Phys. Chem. Chem. Phys.* **2003**, *5*, 2836.

(18) Walter, E.; Pronzato, L. *Identification of Parametric Models from Experimental Data*; Springer-Verlag: Berlin, 1997.

(19) Asai, S.; Nakamura, H.; Sumita, T. Oxidation of Benzyl Alcohol using Hypochlorite Ion Via Phase-Transfer Catalysis. *AIChE J.* **1994**, *40*, 2028.