Model-free multivariate curve resolution combined with model-based kinetics: algorithm and applications

Mathias Sawall[a], Armin Börner[b], Christoph Kubis[b], Detlef Selent[b], Ralf Ludwig[c], Klaus Neymeyr[*a]

[a] *Universität Rostock, Institut für Mathematik, Ulmenstrasse 69, 18057 Rostock, Germany*
[b] *Leibniz-Institut für Katalyse, Albert-Einstein-Strasse 29a, 18059 Rostock*
[c] *Universität Rostock, Institut für Chemie, Dr.-Lorenz-Weg 1, 18059 Rostock, Germany*

Abstract

Multivariate curve resolution techniques are powerful tools to extract from sequences of spectra of a chemical reaction system the number of independent chemical components, their associated spectra and the concentration profiles in time. Usually, these solutions are not unique because of the so-called rotational ambiguity.

In the present work we reduce the non-uniqueness by enforcing the consistency of the computed concentration profiles with a given kinetic model. Traditionally, the kinetic modeling is realized in a separate step which follows the multivariate curve resolution procedure. In contrast to this, we consider a hybrid approach which combines the *model-free* curve resolution technique with the *model-based* kinetic modeling in an overall optimization. For a two-component model problem the range of possible solutions is analyzed and its reduction to a single, unique solution by means of the hybrid kinetic modeling is shown. The algorithm reduces the rotational ambiguity and improves the quality of the kinetic fitting. Numerical results are also presented for a multi-component catalytic reaction system which obeys the Michaelis-Menten kinetics.

Key words: chemometrics, factor analysis, kinetic modeling, pure component decomposition, spectral recovery, hydroformylation.

1. Introduction

Computerized spectroscopic measurements of a chemical reaction system can produce enormous amounts of spectroscopic data. The information content of these data is to be recovered by means of numerical algorithms. Self-modeling curve resolution techniques can be used to detect the number of independent chemical species, their absorption spectra as well as their concentration profiles in time. Sometimes these computed concentration profiles are fitted to a kinetic model in a separate step which follows the factorization process. This approach of an *a-posteriori* kinetic modeling has some disadvantages. The main problem is that self-modeling curve resolution techniques suffer from the so-called rotational ambiguity of the solution. Typically, there is a continuum of mathematically admissible solutions. However, most of these solutions do not contain useful information on the chemical system. A successful strategy to reduce this non-uniqueness is to impose further restrictions on the solutions. e.g., those on the smoothness or on the correlation of the factors. Even then there are still many admissible solutions. However, most of these solutions are not consistent with a pre-given kinetic model. Therefore our aim is to combine a model-free curve resolution technique with a model based kinetic regularization. The regularization approach allows to compute those admissible spectral and concentration factors which are consistent with the kinetic model. This considerably reduces the rotational ambiguity.

Such a combination of a model-free curve resolution technique with a model-based kinetic regularization is not new and has been applied to various problems; see, e.g., the fundamental works of Juan, Maeder, Martínez and Tauler [14], Haario and Taavitsainen [9], and the references therein. However, in [14] the result of a hard-model kinetic fit *replaces* the concentration factor within an MCR-ALS algorithm; this algorithm combines soft- and hard-modeling. In contrast to this we do not replace the concentration factor by its kinetic-model-fit but use the model-error for a regularization of the reconstruction functional, see Section 3.2. Further, the focus of this paper is on an explicit representation

[*] *Correspondence to: K. Neymeyr, Universität Rostock, Institut für Mathematik, Ulmenstrasse 69, 18057 Rostock, Germany.* *June 23, 2012*

of the attainable reduction of the rotational ambiguity due to a kinetic regularization. Numerical results are given for a model problem and for the homogeneously catalyzed hydroformylation process, also known as oxo synthesis.

2. Model-free curve resolution

We consider multivariate data to be given in the form of a time-sequence of spectra of a non-stationary chemical system. As already said, the task of a model-free analysis is to determine the number of independent chemical species, which are often the pure components of the chemical reaction system, and also their concentration profiles in time and the associated absorption spectra. There are several monographs and review papers on this topic, see e.g. [12, 17, 18, 19] and the references therein.

The number of separate spectra is denoted by k. Each spectrum is formed of absorption values given at n separate frequencies. These data are written in the rows of a $k \times n$ data matrix $D \in \mathbb{R}^{k\times n}$. If the chemical reaction system contains a number of s components, i.e., the reactants and products, then Lambert-Beer's law says that D has the form

$$D = CA. \tag{1}$$

Nonlinearities and error terms are ignored in (1). The columns of the matrix $C \in \mathbb{R}^{k\times s}$ are the concentration profiles along the time axis of the components. The rows of the matrix $A \in \mathbb{R}^{s\times n}$ are formed by the spectra of these pure components. For a practical problem the matrix factorization $D = CA$ is not known. The task of a model-free curve resolution technique is to compute the matrix factorization (1) without prior knowledge of the chemical system. The main difficulty is that a plethora of matrix factors C and A exist whose product equals D. In the best case there is only one chemically meaningful solution which is to be determined by means of mathematical tools. Next we consider non-negative solutions, which is a first step to reduce the range of admissible solutions.

2.1. Non-negative matrix factorization

The most restrictive and most useful demand on a factorization $D = CA$ is that only non-negative solutions can be accepted for the concentration factor C and also for the pure component spectra factor A. Negative entries of C and A do not have a physical interpretation. These restrictions give rise to define the following non-negative matrix factorization problem.

Problem 1 (Non-negative matrix factorization)**.** *Given a non-negative matrix $D \in \mathbb{R}_+^{k\times n}$, whose rank equals s, non-negative matrix factors $C \in \mathbb{R}_+^{k\times s}$ and $A \in \mathbb{R}_+^{s\times n}$ are to be determined so that*

$$D = CA.$$

For an arbitrary non-negative matrix D Problem 1 may have no solution. Examples of such matrices without a non-negative factorization have been presented in [4, 28]; one can construct k-by-n matrices D with $2 < \operatorname{rank}(D) < \min(k, n)$ which do not have a non-negative matrix factorization. Fortunately, our situation is somewhat different. The Lambert-Beer law (1) says that D originates from the spectroscopic measurement on a chemical system whose intrinsic components are just specified by C and A. Hence for spectroscopic data D the approximate existence of a non-negative matrix factorization appears to be guaranteed.

The factorization of D into C and A is not unique. It is well known that any solution of the factorization problem is non-unique with respect to the arrangement and positive scaling of the columns of C and the rows of A. This type of non-uniqueness is not really a difficult problem. In contrast to this, there is a further sort of non-uniqueness. Nearly always a continuum of solutions exists which can not be reduced to a single reference solution by a rearrangement and rescaling of the matrix factors C and A. For such problems several analytical methods and numerical methods have been developed [1, 6, 7, 13, 16, 18, 19, 22, 23, 24]. Sometimes it is possible to give a visual representation of the continuum of possible solutions. Such an approach for a three-component system has recently been presented by Golshan, Abdollahi and Maeder [7], see also [2, 3, 20, 25]. The decisive question is to determine the chemically relevant solution within the range of feasible (mathematically admissible) solutions. Usually, a multivariate curve resolution technique does not compute the range of these admissible solutions. Instead *regularization techniques* in the form of so-called *soft constraints* are used to steer the factorization algorithm. In this way one desires to construct the chemically relevant and correct factorization. Such a regularization is also part of the factorization tool PCD (Pure Component Decomposition) which is explained next.

2.2. Factorization tool PCD

Typically the spectral data matrix D includes measurement errors. Further, the Lambert-Beer law (1) is an idealization which neglects any non-linearities. Therefore we aim at an *approximate reconstruction CA* of

D in a way that $D - CA$ is close to the zero matrix. There are other useful restrictions on the matrix factors C and A which are different from the non-negativity conditions. To meet all these conditions we reformulate Problem 1 as a minimization problem for the function

$$F : \mathbb{R}^{k\times s} \times \mathbb{R}^{s\times n} \to \mathbb{R},$$
$$F(C, A) = \|D - CA\|_F^2 + \sum_{i=1}^{p} \gamma_i f_i(C, A). \qquad (2)$$

Therein $\|\cdot\|_F$ denotes the Frobenius matrix norm, which is the square root of the sum of squares of all matrix elements [8]. Further, p is the number of penalty and regularization functions f_i which are weighted by positive constants γ_i whose numerical values are given below. These penalty functions are used to enforce to a large extent the hard constraints on the solution. Further regularization functions with typically smaller weighting constants γ_i compared to the penalty functions can be used to give advantage to solutions which respect the soft constraints. For details on those functions f_i which are used in the PCD code see [23].

Typically a first step to solve the minimization problem is to compute a singular value decomposition (SVD) [8] of the data matrix $D \in \mathbb{R}^{k\times n}$. The SVD reads

$$D = U\Sigma V^T$$

with orthogonal matrices $U \in \mathbb{R}^{k\times k}$ and $V \in \mathbb{R}^{n\times n}$. Further, $\Sigma \in \mathbb{R}^{k\times n}$ is a diagonal rectangular matrix with

$$\Sigma_{i,j} = \begin{cases} \sigma_i, & \text{for } i = j, \\ 0, & \text{else.} \end{cases}$$

The numbers $\sigma_1 \geq \sigma_2 \geq \cdots \geq 0$ are the singular values. A *low-rank approximation* $\tilde{D}$ of D is accessible from the SVD if all singular values which are smaller than a proper threshold are ignored. We obtain $\tilde{D} = \tilde{U}\tilde{\Sigma}\tilde{V}$ with the sub-matrices $\tilde{U} = U(:, 1 : s)$ and $\tilde{V} = V(:, 1 : s)$, which are formed from the s first columns of U and V, and the $s \times s$ submatrix $\tilde{\Sigma} = \Sigma(1 : s, 1 : s)$ of Σ. This low-rank approximation still contains the characteristic information on the chemical system and opens an algorithmic way to solve the optimization problem for the function F.

This approach, which has already been used by Lawton and Sylvestre [16], allows us to approximate F, see (2), in the form

$$G : \mathbb{R}^{s\times s} \to \mathbb{R},\ G(T) = F(\tilde{U}\tilde{\Sigma}T^{-1}, T\tilde{V}^T). \qquad (3)$$

Therein a matrix $T \in \mathbb{R}^{s\times s}$ and its inverse T^{-1} are used. These matrices are the key to represent and to treat the so-called rotational ambiguity of the solution. The next step is a minimization of G with respect to the matrix elements of T. For a proper treatment of perturbed data it may appear advisably to work with a matrix $T \in \mathbb{R}^{s\times z}$ with $z \geq s$; then one has to substitute T^{-1} by its pseudo-inverse T^+ [8]. The use of (T, T^+) admits to get access to additional singular vectors for a proper reconstruction of C and A. The choice of the regularization functions and their weights factors γ_i is important. Negative entries in C or A are penalized by f_1 and f_2. For the problems in Sections 4 and 5 we use $\gamma_1 = \gamma_2 = 100$. Hard constraints penalizing negative entries in C or A are much stronger weighted than soft constraints like the smoothness of the concentration profiles and/or localized, uncorrelated spectra; see also [18, 19, 23].

3. Kinetic modeling

A successful pure component decomposition can be used to extract additional information on the chemical system from the concentration profiles since these functions encode useful information on the kinetics of the chemical reaction. In a subsequent step the kinetic constants of a given kinetic model can be fitted to the concentration profiles. However, the non-uniqueness or rotational ambiguity of the factorization may result in a multitude of concentration profiles which usually do not allow an accurate fitting to a pre-given kinetic model. Therefore one can try to combine the *model-free* factorization approach with a *model-based* chemical kinetic in order to steer the optimization procedure within the set of feasible factorizations to just those solutions which are consistent with the kinetic model [9, 14]. Mathematically this can be realized by adding the squared error of a model fitting as a further regularization term to the function F given in (2).

Typically the kinetics of a chemical reaction can be modeled by a system of ordinary differential equations in the form

$$\frac{d}{dt}c(t) = g(t, c(t), K). \qquad (4)$$

The vector-valued function $c(t)$ represents the concentrations of the chemical species at a time t. If a vector $c(t_1)$ of initial concentrations at a time t_1 is given, then (4) defines an initial value problem. Equation (4) is assumed to hold for all t in an interval $[t_1, t_{\text{end}}]$. The right hand side of (4) is formed by the rate equations of the chemical reaction. The function g also depends on the vector K of the kinetic constants. The problem of a kinetic modeling is to determine just this vector K. How this can be done is explained next.

3.1. A-posteriori kinetic modeling

The "traditional" approach to a kinetic modeling which can be considered as an *a-posteriori* modeling includes two steps. In a first step the concentration profiles C are computed by a pure component decomposition. As mentioned above any decomposition $D = CA$ includes a scaling-ambiguity in the sense that a diagonal matrix Δ with positive diagonal elements can be introduced so that $D = (C\Delta^{-1})(\Delta A)$. Because of this scaling-ambiguity it appears to be possible and convenient to normalize all the absorption spectra, that are the rows of A, in a way that their maximum equals 1. Such a normalization implicitly determines the scaling of the columns of C. Here we use a different normalization, namely, we scale the columns of C in a way that the concentration profiles are consistent with experimental data like initial concentration and mass balances, see Sections 4 and 5. The matrix C with properly scaled columns is denoted by $C^{(\mathrm{S})}$.

In a second step these scaled concentration profiles can be used to compute the kinetic constants K in a way that the solution of (4) is a proper fitting to the concentration profiles $C^{(\mathrm{S})}$. Since $C^{(\mathrm{S})}$ is a $k \times s$ matrix we evaluate the solution $c(t)$ of (4) on the time grid $t_1 < t_2 < \cdots < t_k = t_{\mathrm{end}}$. These t_i are just the points in time which have been used to form the data matrix D. If $C^{(\mathrm{S})}_{i,j}$ are the components of the matrix $C^{(\mathrm{S})}$ and $C^{(\mathrm{ode})}(K) \in \mathbb{R}^{k\times s}$ is the solution of (4) evaluated on the grid $(t_1, \ldots, t_k)$, then

$$f_{\mathrm{kin}} : \mathbb{R}^q \to \mathbb{R},$$
$$f_{\mathrm{kin}}(K) = \sum_{i=1}^{k}\sum_{j=1}^{s}(C^{(\mathrm{S})}_{i,j} - C^{(\mathrm{ode})}_{i,j}(K))^2 \qquad (5)$$

is the squared modeling error for a given vector $K \in \mathbb{R}^q$ of kinetic constants. We denote by K^* the vector of optimal kinetic constants which minimizes f_{kin} so that $C^{(\mathrm{S})}$ and $C^{(\mathrm{ode})}(K^*)$ have the smallest distance (componentwise sum of squares). The vector K^* can be computed by a proper numerical minimization procedure.

3.2. Model-based regularization

The a-posteriori kinetic modeling has some disadvantages. Whenever a continuum of admissible factorizations exists, then a curve resolution technique will often provide a concentration factor C which might not allow a very accurate kinetic fitting with the kinetic model. As explained above it appears to be preferable to apply the error functional (5) as a further penalty function to (3). Then the minimization of f_{kin} is a part of the minimization of the function G. By combining (2), (3) and (5) we obtain

$$\tilde{G}(T, K) = F(\tilde{U}\tilde{\Sigma}T^{-1}, T\tilde{V}^T) + \gamma_{\mathrm{kin}} f_{\mathrm{kin}}(K). \qquad (6)$$

If $\tilde{G}$ attains a minimum in (T^*, K^*), then the process of a kinetic-model-based curve resolution results in $C = \tilde{U}\tilde{\Sigma}(T^*)^{-1}$, $A = T^*\tilde{V}^T$ and a kinetic model (4) with kinetic constants K^*. In other words we use the process of kinetic modeling as a further regularization. The usage of f_{kin} for a regularization appears to be the natural approach; we minimize $\tilde{G}$ in T and K simultaneously. For the minimization we use the NL2SOL code [5] which is based on quasi-Newton and Gauss-Newton algorithms. Alternatively, but we do not use such a strategy, one might minimize $\tilde{G}(T, K)$ in an alternating iteration; then separate minimization steps in T and in K are repeated until convergence in T and in K is achieved.

For the approximate solution of the ordinary differential equation (4) numerical solvers are to be applied. For non-stiff problems we use a high-order explicit Runge-Kutta solver like the Dormand-Prince method with step size control [11]. For stiff problems implicit Runge-Kutta procedures are to be used like the Radau methods [10]. In our PCD code we use Radau IIA methods with the variable orders 5, 9 and 13 together with step-size control[1].

4. A two-component model problem

In order to show that a regularization by means of a kinetic model is a successful strategy we start with an analysis of a two-component model problem. For this problem the continuum of possible solutions is accessible; we show that the hybrid approach using (6) can reduce the ambiguity.

We assume a first order kinetics $X \xrightarrow{K} Y$. The associated concentration profiles for the components X and Y are taken as

$$c_X(t) = \exp(-Kt), \quad c_Y(t) = 1 - \exp(-Kt) \qquad (7)$$

with a kinetic constant $K > 0$. The absorption spectra

[1] The RADAU software in FORTRAN is available from http://www.unige.ch/~hairer/software.html

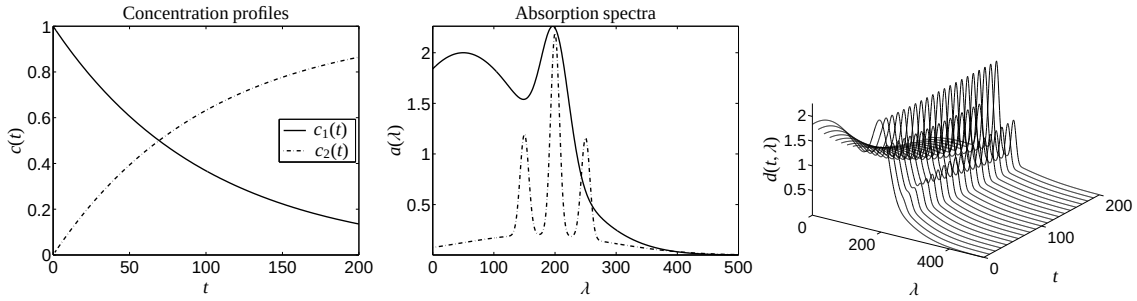

Figure 1: Concentration profiles (left) and absorption spectra (center) of the model problem. Total absorption of the mixture (right).

of the two components X and Y are set to be

$$a_X(\lambda) = 2\exp(-\frac{(\lambda-50)^2}{30000}) + 1.3\exp(-\frac{(\lambda-200)^2}{1000}),$$
$$a_Y(\lambda) = \exp(-\frac{(\lambda-150)^2}{100}) + \frac{1}{5}\exp(-\frac{(\lambda-170)^2}{30000}) + 2\exp(-\frac{(\lambda-200)^2}{100}) + \exp(-\frac{(\lambda-250)^2}{100}).$$

According to Lambert-Beer's law the absorption spectrum of the mixture at a time t and wave-length λ is

$$d(t,\lambda) = c_X(t)a_X(\lambda) + c_Y(t)a_Y(\lambda).$$

The functions $c(t)$, $a(\lambda)$ and $d(t,\lambda)$ with $K = 0.01$ are shown in Fig. 1.

An equidistant grid on $t \in [0, 200]$ and $\lambda \in [0, 500]$ with $k = 201$ and $n = 501$ is used to construct the data matrix D of absorption values. Thus the matrix $D \in \mathbb{R}_+^{201\times501}$ reads

$$D_{i,j} = d(t_i, \lambda_j), \quad i = 1,\ldots,201,\ j = 1,\ldots,501. \quad (8)$$

The data matrix D has the rank 2. Therefore the wanted matrix factors are $C \in \mathbb{R}_+^{201\times2}$ and $A \in \mathbb{R}_+^{2\times501}$. Further, we assume the initial concentrations to be given as $c_X(0) = 1$ and $c_Y(0) = 0$, i.e. the first row of C is the row vector $C(1,:) = (c_X(0), c_Y(0)) = (1, 0)$.

Next the mathematical problem is summarized.

Given:
- $D \in \mathbb{R}_+^{201\times501}$.
- Number of components $s = 2 = \text{rank}(D)$ (by construction).
- Initial concentrations $c_X(0) = 1$, $c_Y(0) = 0$.

Aim:
- A non-negative decomposition $D = CA$.
- The kinetic parameter K.

4.1. Continuum of admissible factorizations

Next we present an analytic representation of the range of feasible non-negative factorizations of the data matrix (8). The starting point is a singular value decomposition $U\Sigma V^T$ of D. As $\text{rank}(D) = 2$ only those left and right singular vectors are needed which correspond to non-zero singular values. Therefore let $\tilde{U} = U(:, 1:2)$, $\tilde{\Sigma} = \Sigma(1:2, 1:2)$ and $\tilde{V} = V(:, 1:2)$ so that $D = \tilde{U}\tilde{\Sigma}\tilde{V}^T$ if numerical rounding errors are ignored. This factorization of D allows to represent the range of admissible non-negative matrix factorizations. By inserting a regular transformation matrix $T \in \mathbb{R}^{2\times2}$ and its inverse we obtain

$$C = \tilde{U}\tilde{\Sigma}T^{-1},\ A = T\tilde{V}^T.$$

Only those regular T are considered which result in non-negative matrices C and A.

The assumption $c_X(0) = 1$ determines the unknown calibration constant of the first column of C; the unknown scaling constant of the second column of C can be determined in the least-squares sense from the constraint that $c_X(t) + c_Y(t) = c_X(0) = 1$ for all t. Having fixed the row of C as $(1, 0)$ the wanted pure component spectrum of X is just the first row $D(1,:)$ of D, i.e. $A(1,:) = D(1,:)$ since at $t = 0$ only the component X is present. Therefore the first row of the transformation matrix T is determined; the real constants α and β in

$$T(x) = \begin{pmatrix} \alpha & \beta \\ 1 & x \end{pmatrix}$$

can be computed from $A(1,:) = D(1,:) = T(1,:)\tilde{V}^T$. As the second spectrum $A(2,:)$ can be scaled in any (non-zero) way, the second row of T can be set to $(1, x)$; Golshan, Abdollahi and Maeder use the same approach in Eq. (7) of [7]. Thus $x \in \mathbb{R}$ is the only remaining degree of freedom.

We note that the use of $(1, x)$, instead of working with $(-1, x)$, is justified by the Perron-Frobenius theorem [21]. This theorem allows to assume that the first singular vectors $U(:, 1)$ and $V(:, 1)$ which correspond to

the largest singular value of the non-negative data matrix D are non-negative. Hence, $\min_i V_{i,1} \geq 0$. Further, the orthogonality of $V(:,1)$ and $V(:,2)$ guarantees that $\min_i V_{i,2} < 0 < \max_i V_{:,2}$. Therefore a positive contribution from $V(:,1)$ is required in order to construct a non-negative matrix A. See also [16] for the use of the Perron-Frobenius theorem.

This results in a one-parameter representation of the absorption and concentration matrices which reads

$$A[x] = \begin{pmatrix} \alpha & \beta \\ 1 & x \end{pmatrix} \tilde{V}^T,$$
$$C[x] = \tilde{U}\tilde{\Sigma}\frac{1}{\alpha x - \beta}\begin{pmatrix} x & -\beta \\ -1 & \alpha \end{pmatrix}.$$

To guarantee non-negativeness of the elements in the second row $A(2,:)[x]$ the parameter x has to satisfy

$$\max_{\substack{i=1,\ldots,n,\\ \tilde{V}_{i,2}>0}} -\frac{\tilde{V}_{i,1}}{\tilde{V}_{i,2}} \leq x \leq \min_{\substack{i=1,\ldots,n,\\ \tilde{V}_{i,2}<0}} -\frac{\tilde{V}_{i,1}}{\tilde{V}_{i,2}}. \tag{9}$$

The non-negativeness of the elements of $C(x)$ results in the following restrictions

$$x > \frac{\beta}{\alpha} \text{ and } x \geq \max_i \frac{\sigma_2 \tilde{U}_{i,2}}{\sigma_1 \tilde{U}_{i,1}} \text{ if } \tilde{U}\tilde{\Sigma}(-\beta,\alpha)^T \geq 0, \tag{10}$$

$$x < \frac{\beta}{\alpha} \text{ and } x \leq \min_i \frac{\sigma_2 \tilde{U}_{i,2}}{\sigma_1 \tilde{U}_{i,1}} \text{ if } \tilde{U}\tilde{\Sigma}(-\beta,\alpha)^T \leq 0. \tag{11}$$

All these restrictions cannot reduce the range of admissible solutions to a single, unique solution. Instead for the whole range $x \in [0.4286, 0.9405]$ non-negative solutions exist. The possible solutions $C(:,1)[x]$ and $A(2,:)[x]$ are shown in Fig. 2.

4.2. Regularization with a kinetic model

Next we follow the strategy of a model-based regularization as explained in Sec. 3.2. The kinetic model (7) is taken as a part of the penalty function, see Eqns. (5) and (6). This defines in a non-linear least-squares problem for the two parameters x and K. A numerical computation for the given test problem shows that these parameters are confined as follows

$$(x, K) \in [0.4286, 0.9405] \times (0, \infty).$$

For each $x \in [0.4286, 0.9405]$ we obtain an admissible solution of the factorization problem - however for most of the x this solution does not reproduce the original spectra and concentration profiles.

In order to reconstruct the correct solution we use an error functional measuring the error between a matrix

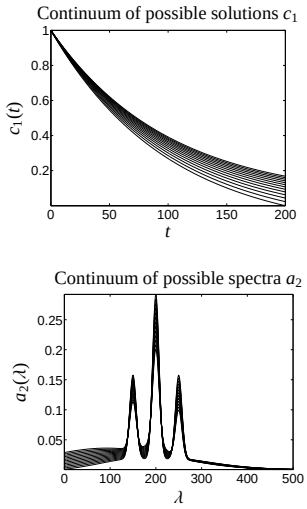

Figure 2: Continua of solutions for $c_1 = C(:,1)[x]$ and $a_2 = A(2,:)[x]$.

factor $C \in \mathbb{R}^{201\times 2}$ and the solution $C^{(\text{ode})}$ of (4). For the model problem the solution of (4) is explicitly given in (7) so that f_{kin} reads

$$f_{\text{kin}} : \mathbb{R}^2 \to \mathbb{R},\ (x, K) \mapsto \sum_{i=1}^{400} (g_i(x, K))^2. \tag{12}$$

This function is to be minimized. Therein the g_i are

$$g_i(x, K) = C_{i,1}[x] - C_{i,1}^{(\text{ode})}(K), \qquad i = 1, \ldots, 200,$$
$$g_{i+200}(x, K) = C_{i,2}[x] - C_{i,2}^{(\text{ode})}(K), \quad i = 1, \ldots, 200.$$

Note that the concentration profile $C(:,2)$ also depends on x since the transformation T couples the scaling of $C(:,2)$ to that of $C(:,1)[x]$.

Fig. 3 shows the regularization functional f_{kin} for $K \in (0, 0.025]$. To illustrate that arbitrary $x \in [0.4286, 0.9405]$ can result in concentration profiles which are not consistent with the kinetic model we take arbitrarily $x = 0.5$ and compute the optimal kinetic fitting which results in $K = 0.012382$. For this solution the hard constraints, namely the non-negativity of C and A, are satisfied; the error on these parts of the penalty functions is less that 10^{-10}. However, the concentration profiles are only poor approximations of (7) where $K = 0.01$. The kinetic-model-error $f_{\text{kin}}(K)$, as given by (5), is about 0.2, see Fig. 3. Next we show that the kinetic regularization is the key for a correct reconstruction of C and A.

Figure 3 shows that a kinetic-model-error close to 0 can be achieved. And in fact, the hybrid algorithm using the kinetic regularization (12) can provide the correct solution. The numerical computation yields $x^* =$

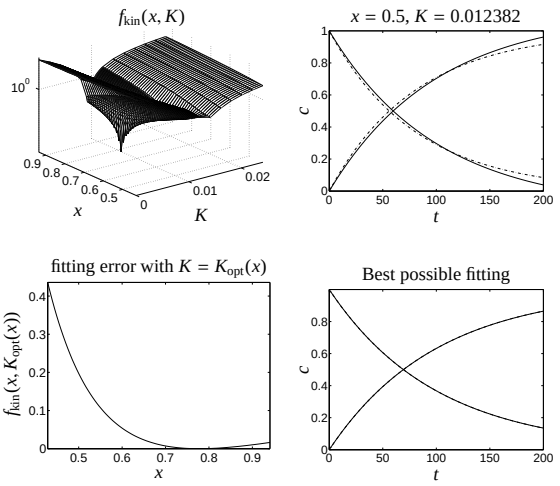

Figure 3: Upper left: The function $f_{kin}(x, K)$; Upper right: The exact decomposition with optimized K for $x = 0.5$ results in a poor fitting. Lower left: The kinetic fitting error $f_{kin}(x, K_{opt})$ with optimal $K_{opt} = K_{opt}(x)$ as a function of x. Minimum in $x^* = 0.77753$. Lower right: The concentration profiles for the globally optimal x^* and K^* which reproduces the original $c(t)$ shown in Fig. 1.

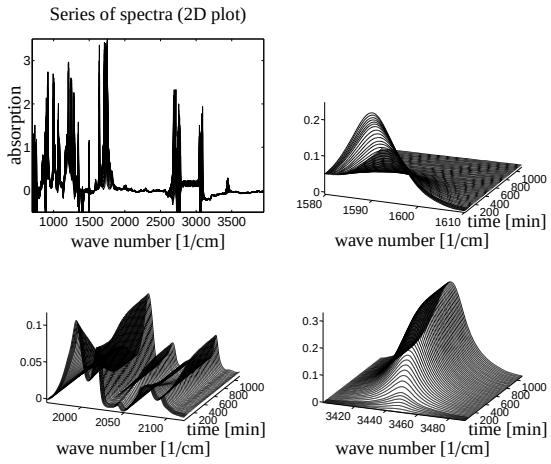

Figure 4: Upper left: Complete series of spectra after background subtraction; Further plots: Characteristic bands after zero-line correction.

0.77753 together with $K^* = 0.0099973$ which is a very good approximation of the desired solution, see Fig. 3.

For the artificial two-component model problem the model-based regularization has proved as a well-working tool to steer the curve resolution method to the correct solution.

5. Rhodium-catalyzed hydroformylation

The rhodium catalyzed hydroformylation of 3,3-dimethyl-1-butene, forming 4,4- dimethylpentanal and 2,3,3-trimethylbutanal in a constant 9:1 ratio, has been studied over the full conversion range. The reaction was performed in n-hexane solvent at 30°C, with p(CO) = 1 MPa and p(H_2) = 2 MPa. The catalyst, a hydrido rhodium carbonyl phosphite complex, was formed from [(acac)Rh(CO)$_2$] and tri(2,4-di-tert.-butylphenyl)phosphite prior to the catalytic reaction. The individual rhodium, phosphite and olefin concentration applied were $3 \cdot 10^{-4}$, $6 \cdot 10^{-3}$ and 0.9 mol dm^{-3}, respectively. Catalyst preformation, the progress of the organic reaction as well as catalytic intermediates have been monitored by in situ FTIR-spectroscopy. For that purpose, the reaction solution was circulated between the batch reactor and a pressure tight transmission cell with the windows material ZnS placed inside a Bruker Tensor 27 FTIR spectrometer. The background spectrum used for correction consisted of the solvent n-hexane and dodecane, with the latter serving as an internal standard for GC. The independent concentration profiles of the product aldehydes obtained from GC analysis were found to fit well to the data derived from FTIR spectroscopy. To get access to the concentration profiles of the catalytically relevant rhodium complexes, appropriate spectral data were selected and treated by the PCD algorithm.

For this problem the data matrix is built from 1045 spectra, each with 13482 spectral channels. The curve resolution algorithm makes use only from three decisive spectral bands namely $[1580, 1601]\text{cm}^{-1}$, $[1960, 2120]\text{cm}^{-1}$ and $[3400, 3490]\text{cm}^{-1}$. Further, the background spectrum, i.e., the spectrum of the solvent and n-hexane, is subtracted and the zero-line is corrected. This results in a reduced matrix $D \in \mathbb{R}^{1045\times1170}$. The complete series of spectra after zero-line correction as well as the three isolated spectral bands are shown in Fig. 4.

The singular value decomposition of D provides valuable information on the number of independent species. The first eight left and eight right singular vectors and also the twenty largest singular values are shown in Fig. 5. There are four smooth or non-oscillatory left singular vectors and at least four singular values which are clearly separated from the remaining set of smaller singular values. The noise pattern of the left singular vectors with the indexes $5, \ldots, 8$ does no appear for the corresponding right singular but occurs for larger indexes. Our explanation is that this caused by the background subtraction and the zero-line correction, which each operate along the frequency axis.

The system contains at least four independent reactants which determines the dimensions of the matrix factors $C \in \mathbb{R}_+^{1045\times4}$ and $A \in \mathbb{R}_+^{4\times1170}$. The curve res-

olution technique should compute these factors so that $D \approx CA$. Further, we assume a Michaelis-Menten kinetics [15]. The vector of kinetic constants K is to be computed in a way that the solution of the ordinary differential equation fits the concentration profiles C from the pure component decomposition in the least-squares sense.

In Sec. 5.1 a non-negative matrix factorization of D is used to compute C and A without any kinetic modeling. The resulting mathematically admissible factorization is not consistent with the Michaelis-Menten kinetics. The *model-based regularization* in Sec. 5.2 is the key to compute proper pure component spectra together with concentration profiles which are compatible with the Michaelis-Menten kinetics.

5.1. Results of an SMCR analysis

A first approach to solve the chemometric problem for the Rhodium-catalyzed hydroformylation is to compute a non-negative matrix factorization as defined by Problem 1 in Sec. 2.1. For the moment we ignore any kinetic modeling. However, two soft-constraints are used. First a norm of the discrete second derivative of the concentration profiles is used as a penalty function in order to give an advantage to smooth concentration profiles. Further, the discrete integral of the spectra is taken for a regularization in order to favor localized peaks in the spectrum. This soft-constrained non-negative matrix factorization can, e.g., be computed by the PCD algorithm [23]. The weighting factor for the discrete-integral-regularization is $\gamma = 1 \cdot 10^{-4}$ whereas the weighting factors for the penalization of negative entries in C and A are equal to 100.

As explained in Sec. 2.2, see also [23], the truncated singular value decomposition is used to reconstruct the matrix factors by using six left and right singular vectors, i.e., $s = 4$ and $z = 6$. The results are shown in Fig. 6. Therein the absorption spectra are normalized in a way that the maximum is set to 1. This implicitly determines the associated scaling of the concentration profiles.

5.1.1. Non-uniqueness - a continuum of decompositions

It is easy to see that there is a continuum of non-negative solutions. One of these solutions is shown in Fig. 6. To the end of an illustration of a subset of this continuum we construct a one-dimensional range of these admissible solutions. This range is parameterized next by a single parameter α. For the explicit construction we apply the following transformation to the matrix

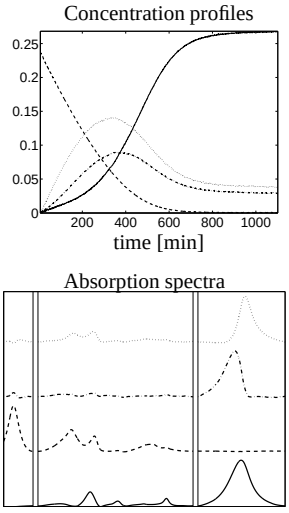

Figure 6: An admissible non-negative decomposition with a small reconstruction error, smooth concentration profiles and localized peaks so that the integral of the absorption spectra is small.

factors

$$C' = CT_\alpha, \quad A' = T_\alpha^{-1}A.$$

Therein T_α is taken as a 4-by-4 matrix

$$T_\alpha = I_4 - \alpha e_4 e_3^T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -\alpha & 1 \end{pmatrix}, \qquad (13)$$

where I_4 is the 4×4 identity matrix and e_k is its k-th column. If $\alpha \in [0, 0.3127]$, then C' and A' are still non-negative matrices, which is easy to see by checking the associated linear combinations of the columns $C(:,3)$ and $C(:,4)$ of C. Further, T_α^{-1} is a non-negative matrix so that $A' = T_\alpha^{-1}A$ is also non-negative.

There are many further transformations which preserve the non-negativity of the matrix factors like $T_\beta = I_4 - \beta e_3 e_1^T$ with $\beta \in [0, 0.4096]$ or $T_\delta = I_4 - \delta e_4 e_1^T$ with $\delta \in [0, 0.16]$. Various other combinations or products of these factors may also work. Fig. 7 illustrates the ambiguity of the solution due to the transformation (13).

The so-called rotational ambiguity of the solution depends to some extent on the dimension s. In case of multi-component systems with $s \geq 2$ the knowledge of a specific concentration profile does not determine the associated absorption spectrum and vice versa. However, there are some special cases in which more information is available; see [26, 27].

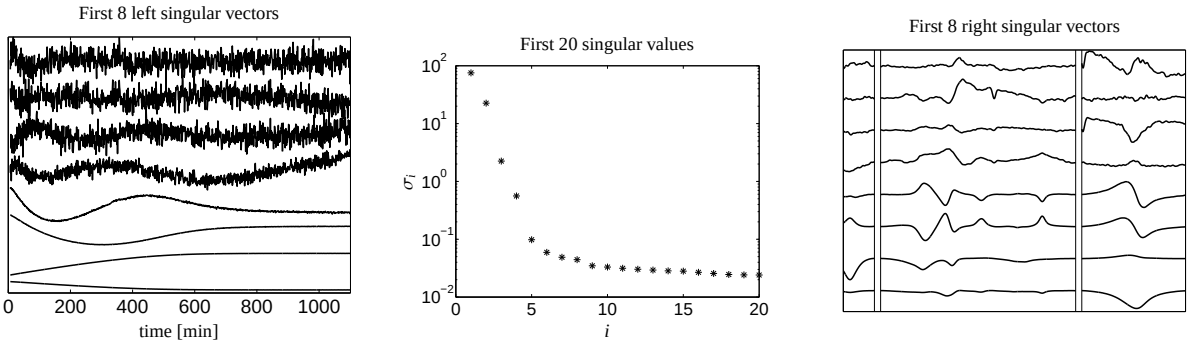

Figure 5: Singular value decomposition of D. Left: Eight left singular vectors corresponding to the largest singular values. Center: Plot of the largest twenty singular values in a semi-logarithmic plot. Right: Eight right singular vectors corresponding to the largest singular values.

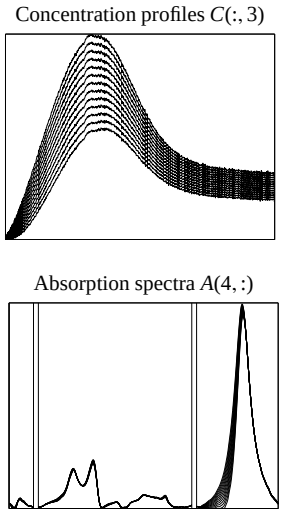

Figure 7: Continuum of admissible non-negative solutions $C(:,3)$ and $A(4,:)$ due to the transformation (13) applied to Fig. 6.

5.1.2. Non-uniqueness - the subsystem of catalysts

The intention of this section is to show that even the subsystem of catalysts contains some inherent rotational ambiguity. The spectra of this subsystem have been gained by a subtraction of the reactant and product spectra from the original spectral data. Next we try to present this ambiguity in detail which is a somewhat technical procedure. Later in Section 5.2 the full four-component problem is treated once again. Then the kinetic regularization technique proves its benefits as everything becomes very simple and the ambiguity disappears.

In order to separate the catalytic subsystem we first state that the absorption spectra are known for the reactant, namely the alkene, and the product being the aldehyde. By using some isolated peaks of the alkene and aldehyde we also obtain their concentration profiles along the time axis. This allows us to subtract the contributions of the reactant and product from the spectral data.

For the resulting catalytic subsystem which contains the acyl- and the hydrido-complex a full uniqueness of the decomposition cannot be attained. There are two further ways to reduce the rotational ambiguity of the system. First, we can exploit the fact that the concentration of a specific component is zero at the end of the reaction. Second, the rotational ambiguity can be reduced if the series of spectra contains a specific spectral band in which only one component is absorbent while all other components show a small absorptivity.

For our catalytic system the acyl complex disappears at the end of the reaction and further an isolated peak in the acyl-complex spectrum is located around 1996cm^{-1} where the absorption of the hydrido complex is close to zero. All this allows to reduce the ambiguity to a one-parametric continuum of admissible solutions. We observe that the concentration profile of the acyl-complex and the absorption spectrum of the hydrido complex appear to be unique apart from small spectral perturbations. However, the complementary concentration profile of the hydrido-complex and the absorption spectrum of the acyl-complex show a *one-parametric continuum* of admissible solutions. All these factors are shown in Fig. 8.

Such an ambiguity makes an *a-posteriori* kinetic fitting difficult. However, a model-based regularization appears to be a promising alternative to compute a useful factorization and reliable kinetic constants as is shown in the next section.

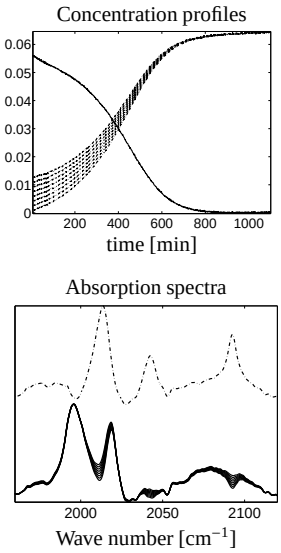

Figure 8: Continuum of solutions of the catalytic subsystem built from the acyl- and hydrido-complex. Uniqueness of the concentration profile of the acyl-complex and the spectrum of the hydrido complex. Ambiguity in the form of a one-parameter continuum for the concentration profile of the hydrido-complex and the spectrum of the acyl-complex.

5.2. Regularization be means of a kinetic model

To approximate the catalytic process we assume a Michaelis-Menten model

$$A + B \underset{K_{-1}}{\overset{K_1}{\rightleftharpoons}} C \xrightarrow{K_2} D + B.$$

The olefin A forms with the catalyst B the catalyst-substrate adduct C. Then C decays into the aldehyde D and catalyst B. The kinetic constants are $K = (K_1, K_{-1}, K_2)$. For a kinetic analysis of the hydroformylation process via in-situ HP-IR and HP-NMR spectroscopy see [15] and the references therein.

The concentration functions are the components of the vector

$$c(t) = \begin{pmatrix} c_A(t) \\ c_B(t) \\ c_C(t) \\ c_D(t) \end{pmatrix}$$

and the system of ordinary differential equations reads

$$\frac{d}{dt}\, c(t) = \begin{pmatrix} -K_1\, c_A\, c_B + K_{-1}\, c_C \\ -K_1\, c_A\, c_B + K_{-1}\, c_C + K_2\, c_C \\ K_1\, c_A\, c_B - K_{-1}\, c_C - K_2\, c_C \\ K_2\, c_C \end{pmatrix}. \qquad (14)$$

As explained in Sec. 4.2 the error of the kinetic modeling is used as a regularization function.

The problem is to minimize the function $\tilde{G}(T, K)$ as given by (6). The result of the minimization are the optimal kinetic constants K^* together with the transformation matrix T - the latter matrix determines the matrix factors C and A.

Within the optimization procedure a proper scaling of the concentration profiles, namely the columns of C, has to be determined in order to evaluate the kinetic-error-functional (15). This is done as follows: For the hydroformylation the mass balance of the organic species, that are the olefin and the aldehyde, says that $C(:,1)+C(:,4)$ equals component-wise 0.8951. Further, the mass balance of the Rhodium-containing species says that $C(:,2) + C(:,3)$ equals in each component $2.9330 \cdot 10^{-4}$. These two equations allow to compute the four scaling constants of the four concentration profiles $C(:,i)$, $i = 1, \ldots, 4$, in such a way that the properly scaled concentration profiles fulfill the mass balance equations in the least-squares sense in the best way. We denote the resulting scaled concentration matrix by $C^{(\mathrm{S})} \in \mathbb{R}^{k\times 4}$.

For the numerical minimization of (6) an implicit Runge-Kutta method Radau IIA with the variable orders (5, 9, 13) has been used. The initial concentrations are $c(0) = (0.8951,\ 2.9330{\cdot}10^{-4},\ 0,\ 0)^T$. This computation works with $C^{(\mathrm{ode})} \in \mathbb{R}^{k\times 4}$ and with the regularization function

$$f_{\mathrm{kin}}(C, K) = \sum_{j=1}^{4} \|C^{(\mathrm{S})}(:,j)\|_\infty^{-1} \sum_{i=2}^{k} (C_{i,j}^{(\mathrm{S})} - C_{i,j}^{(\mathrm{ode})})^2. \quad (15)$$

For $f_{\mathrm{kin}}(C, K)$ we used $\gamma_{\mathrm{kin}} = 100$ as the weighting factor in $\tilde{G}(T, K)$. The factors $\|C^{(\mathrm{S})}(:,j)\|_\infty^{-1}$ in (15) are introduced to supply each of the concentration profiles with the same weighting factor; $\|\cdot\|_\infty$ denotes the maximum norm of a vector [8].

5.3. The numerical results

The final numerical results of the pure component decomposition with embedded kinetic regularization is shown in the Figs. 9 and 10. The absorption spectra can clearly be interpreted; for the details see [15]. In [15] the peaks of the acyl-complex at 2072cm^{-1} and at 2079cm^{-1} are somewhat better separated. Further, the lowly concentrated catalytic species do not have an absorption peak at 1590cm^{-1}. This is just an artifact from the olefin.

The optimization procedure yields the kinetic constants

$$\begin{aligned} K^* &= (K_1^*, K_{-1}^*, K_2^*)^T \\ &= (45.228\,\mathrm{l\,min^{-1}mol^{-1}},\ 0.3431\,\mathrm{min^{-1}}, \\ &\quad 9.3145\,\mathrm{min^{-1}})^T. \end{aligned}$$

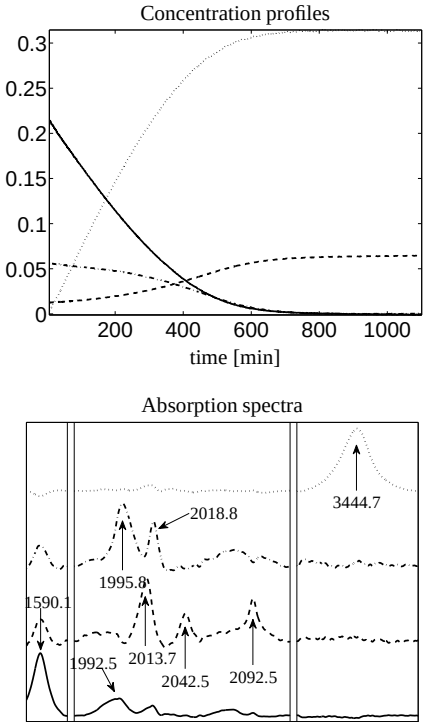

Figure 9: Results of the pure component decomposition by PCD. The maximum of the absorption spectra is scaled to 1 which determines the scaling factors of the associated concentration profiles. Solid line: Olefin. Broken line: Hydrido-complex. Dash-dotted line: Acyl-complex. Dotted line: Aldehyde.

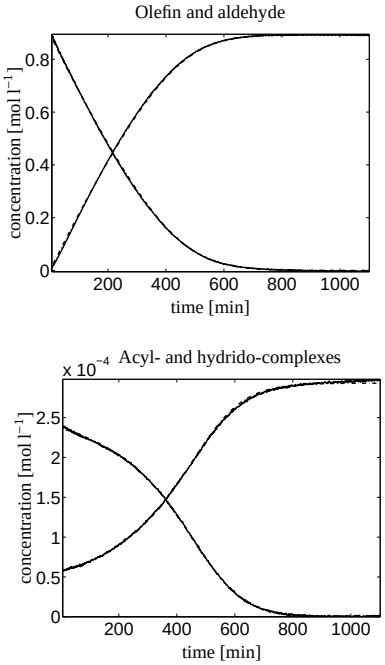

Figure 10: Nearly perfect kinetic fitting of the solution of the Michaelis-Menten equation (14) with the optimal kinetic constants K^* (solid lines) and of the concentration profiles from the pure component decomposition with model-based regularization (dotted lines).

The solution of the Michaelis-Menten system of ordinary differential equations (14) with these constants accurately fits the results of the pure component decomposition, see Fig. 10. The final errors for the four components $(A, B, C, D) = (1, 2, 3, 4)$ are

$$\varepsilon_i = \frac{\|C^{(\mathrm{S})}(:,i) - C^{(\mathrm{ode})}(:,i)\|_2}{\|C^{(\mathrm{S})}(:,i)\|_2}, \quad i = 1, \ldots, 4,$$

$$\varepsilon = (6.15\cdot10^{-3},\ 9.12\cdot10^{-3},\ 7.16\cdot10^{-3},\ 7.67\cdot10^{-3}).$$

Therein the Euclidean norm $\|\cdot\|_2$ is the square root of the sum of the squares of the components of the error.

We note that the computation of the kinetic equilibrium constants K_1^* and K_{-1}^* is poorly conditioned. If we consider only the dependence of the regularizing function f_{kin} on the kinetic constants K_1^*, K_{-1}^* and K_2^* we obtain for its gradient

$$\nabla f_{\mathrm{kin}}(K^*)^T = \begin{pmatrix} 2.9350\cdot10^{-6}, \\ 4.1869\cdot10^{-6}, \\ -3.6041\cdot10^{-6} \end{pmatrix}.$$

So the first order corrections in a Taylor expansion are small. The second-order correction are determined by the Hessian $\nabla^2 f_{\mathrm{kin}}(K^*)$. Its eigenvalues are $\lambda_1 = 2.195\cdot10^{-4}$, $\lambda_2 = 1.2513$ and $\lambda_3 = 11.071$. The second order term $(v, \nabla^2 f_{\mathrm{kin}}(K^*)v)$ of the Taylor expansion is more or less insensitive with respect to variations in the direction of the eigenvector $v_1 = (9.7798\cdot10^{-1}, 2.0870\cdot10^{-1}, -4.1761\cdot10^{-5})^T$ corresponding to the smallest eigenvalue λ_1. Further, v_1 has significant components in the first and second components which are associated with K_1 and K_{-1}. Therefore the numerical computation of these kinetic constants appears to be somewhat instable.

We further note that in Fig. 9 the formation of the acyl-complex is not represented. The formation of this complex is a fast reaction which proceeds in about the first 6 seconds. However, it is possible to deduce the behavior of the reaction system within the first minute without spectral data by extrapolation [15].

6. Conclusion

Model-free multivariate curve resolution techniques suffer from the rotational ambiguity of the solution. This non-uniqueness considerably increases with the number of chemical components. Therefore a considerable proportion of the computed factors of a multi-

component system do not allow a useful chemical interpretation. However, multi-component chemical systems like catalytic reaction systems and their detailed mechanistic understanding are of major importance.

Any additional constraints or restrictions on the solution can help to reduce the non-uniqueness. A very useful supplemental information is that on an underlying kinetic model and even an incomplete kinetic model can be used. If the consistency in the sense of a small fitting error of a kinetic model with the concentration profiles is used to regularize the reconstruction functional, then this ambiguity can drastically be reduced. The resulting numerical algorithm merges the model-free curve resolution technique with the model-based kinetics within an overall optimization process. In the best case reliable and unique multi-component factorizations can be computed.

References

[1] H. Abdollahi, M. Maeder, and R. Tauler. Calculation and Meaning of Feasible Band Boundaries in Multivariate Curve Resolution of a Two-Component System. *Analytical Chemistry*, 81(6):2115–2122, 2009.

[2] O.S. Borgen, N. Davidsen, Z. Mingyang, and Ø. Øyen. The multivariate N-Component resolution problem with minimum assumptions. *Microchimica Acta*, 89:63–73, 1986. 10.1007/BF01207309.

[3] O.S. Borgen and B.R. Kowalski. An extension of the multivariate component-resolution method to three components. *Analytica Chimica Acta*, 174(0):1–26, 1985.

[4] J. E. Cohen and U. G. Rothblum. Nonnegative ranks, decompositions, and factorizations of nonnegative matrices. *Linear Algebra and its Applications*, 190:149–168, 1993.

[5] J.E. Dennis, D.M. Gay, and R.E. Welsch. NL2SOL–An adaptive nonlinear least squares algorithm. *ACM Transactions on Mathematical Software*, 7(3):369–383, 1981.

[6] P. J. Gemperline. Computation of the Range of Feasible Solutions in Self-Modeling Curve Resolution Algorithms. *Analytical Chemistry*, 71(23):5398–5404, 1999.

[7] A. Golshan, H. Abdollahi, and M. Maeder. Resolution of Rotational Ambiguity for Three-Component Systems. *Analytical Chemistry*, 83(3):836–841, 2011.

[8] G.H. Golub and C.F. Van Loan. *Matrix computations*. Johns Hopkins University Press, Baltimore, MD, third edition, 1996.

[9] H. Haario and V.M. Taavitsainen. Combining soft and hard modelling in chemical kinetics. *Chemometr. Intell. Lab.*, 44:77–98, 1998.

[10] E. Hairer and G. Wanner. *Solving ordinary differential equations II, 2nd edition*. Springer, 2002.

[11] E. Hairer, G. Wanner, and S. P. Nørsett. *Solving ordinary differential equations I, 2nd edition*. Springer, 2002.

[12] J.C. Hamilton and P.J. Gemperline. Mixture analysis using factor analysis. II: Self-modeling curve resolution. *Journal of Chemometrics*, 4:1–13, 1990.

[13] J. Jaumot and R. Tauler. MCR-BANDS: A user friendly MATLAB program for the evaluation of rotation ambiguities in Multivariate Curve Resolution. *Chemometrics and Intelligent Laboratory Systems*, 103(2):96–107, 2010.

[14] A. Juan, M. Maeder, M. Martínez, and R. Tauler. Combining hard and soft-modelling to solve kinetic problems. *Chemometr. Intell. Lab.*, 54:123–141, 2000.

[15] C. Kubis, D. Selent, M. Sawall, R. Ludwig, K. Neymeyr, W. Baumann, R. Franke, and A. Börner. Exploring between the extremes: Conversion dependent kinetics of phosphite-modified hydroformylation catalysis. Accepted for Chemistry - A European Journal, 2012.

[16] W. H. Lawton and E. A. Sylvestre. Self modelling curve resolution. *Technometrics*, 13:617–633, 1971.

[17] Y.-Z. Liang, O.M. Kvalheim, and R. Manne. White, grey and black multicomponent systems: A classification of mixture problems and methods for their quantitative analysis. *Chemometrics and Intelligent Laboratory Systems*, 18(3):235 – 250, 1993.

[18] M. Maeder and Y.M. Neuhold. *Practical data analysis in chemistry*. Elsevier, Amsterdam, 2007.

[19] E. Malinowski. *Factor analysis in chemistry*. Wiley, New York, 2002.

[20] E. R. Malinowski. Window factor analysis: Theoretical derivation and application to flow injection analysis data. *Journal of Chemometrics*, 6(1):29–40, 1992.

[21] H. Minc. *Nonnegative matrices*. John Wiley & Sons, New York, 1988.

[22] S. Navea, A. de Juan, and R. Tauler. Detection and Resolution of Intermediate Species in Protein Folding Processes Using Fluorescence and Circular Dichroism Spectroscopies and Multivariate Curve Resolution. *Analytical Chemistry*, 74(23):6031–6039, 2002.

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

[24] R. Rajkó. Computation of the range (band boundaries) of feasible solutions and measure of the rotational ambiguity in self-modeling/multivariate curve resolution. *Analytica Chimica Acta*, 645(1–2):18–24, 2009.

[25] R. Rajkó. Some surprising properties of multivariate curve resolution-alternating least squares (MCR-ALS) algorithms. *Journal of Chemometrics*, 23(4):172–178, 2009.

[26] M. Sawall. Regularisierte nichtnegative Matrixfaktorisierungen und ihre Anwendungen in der Spektroskopie. PhD thesis, Universität Rostock, 2011.

[27] M. Sawall, C. Fischer, D. Heller, and K. Neymeyr. Reduction of the rotational ambiguity of curve resolution technqiues under partial knowledge of the factors. Complementarity and coupling theorems. Technical Report, Universität Rostock and LIKAT, accepted for Journal of Chemometrics, 2012.

[28] L. B. Thomas. Rank factorization of nonnegative matrices (A. Berman). *SIAM Review*, 16(3):393–394, 1974.