

## Estimating reaction rate constants from a two-step reaction: a comparison between two-way and three-way methods

Sabina Bijlsma and Age K. Smilde\*

*Process Analysis and Chemometrics, Department of Chemical Engineering, University of Amsterdam, Nieuwe Achtergracht 166, NL-1018 WV Amsterdam, The Netherlands*

### SUMMARY

In this paper, two different spectral datasets are used in order to estimate reaction rate constants using different algorithms. Dataset 1 consists of short-wavelength near-infrared (SW-NIR) spectra taken in time of the two-step epoxidation of 2,5-di-*tert*-butyl-1,4-benzoquinone using *tert*-butyl hydroperoxide and Triton B catalyst. This dataset showed moderate reproducibility. Dataset 2 consists of UV-VIS recorded spectra of the consecutive reaction of 3-chlorophenylhydrazonopropane dinitrile with 2-mercaptoethanol. This dataset showed good reproducibility. Two-way and three-way methods were used in order to estimate the reaction rate constants for both datasets. For the SW-NIR dataset the lowest standard deviations for the reaction rate constants were obtained with a two-way method. The lowest standard deviations for the reaction rate constant estimates for the UV-VIS dataset were obtained with a two-way method which uses spectral information that is known in advance. In this case the pure spectrum of two reacting absorbing species is known in advance and this information was used by the two-way method. For one two-way method and a few three-way methods which do not use spectral information that is known in advance, pure spectra of the reacting absorbing species of the UV-VIS dataset were estimated which showed excellent agreement with the recorded pure spectra. The pure spectra of the reacting absorbing species for the SW-NIR dataset were not estimated, because it was not possible to record the real pure spectra of these species. For both spectral datasets, quality assessment has been performed using a jackknife method. Copyright © 2000 John Wiley & Sons, Ltd.

KEY WORDS: curve resolution; GRAM; PARAFAC; UV-VIS; SW-NIR

### 1. INTRODUCTION

Kinetics is an important field of interest in chemistry [1–3]. In the literature there are many methods available in order to estimate reaction rate constants from chemical reactions [4–20]. The simplest method to estimate reaction rate constants is curve fitting of kinetic expressions to concentration versus time data [4–6]. However, there are a number of drawbacks in using this approach. It is time-consuming and extraordinarily expensive, since concentration data of the reaction of interest have to be available.

Curve resolution methods adapted to estimate reaction rate constants from spectral data have

---

\* Correspondence to: A. K. Smilde, Process Analysis and Chemometrics, Department of Chemical Engineering, University of Amsterdam, Nieuwe Achtergracht 166, NL-1018 WV Amsterdam, The Netherlands.  
E-mail: asmilde@its.chem.uva.nl

become very popular. Curve resolution is based on the recovery of response profiles and the determination of qualitative information. Parameters of interest, e.g. reaction rate constants, can be incorporated as unknowns [7,21]. Many modifications of different curve resolution-based methods have been published [22–28].

Methods based on three-way analysis to estimate reaction rate constants from batch processes are quite new [15–17]. The principle of these methods is simple. The first method, originally developed by Windig and Antalek [29–32] in order to estimate specific parameters from nuclear magnetic resonance (NMR) data, is a modification of the generalized rank annihilation method (GRAM) [33]. In this modified GRAM a single dataset (master dataset) of a reacting system is split into two datasets (slave datasets) by means of a time shift such that there is a constant time lag between the two slave datasets. Owing to the properties of exponential functions, there exists a special trilinear structure if the slave datasets are stacked. From the decomposition which can be achieved by solving a generalized eigenvalue problem (GEP) [30], specific parameters of interest can be estimated. GRAM is only applicable if the contribution of different species in the mixture spectra is of exponential nature. Hence it is possible to apply GRAM in order to estimate reaction rate constants from spectral data of the reacting system, because kinetic rate equations have an exponential nature. A very important property of GRAM is that in kinetics the method can only be used for (pseudo-)first-order problems.

GRAM is a non-iterative algorithm, which is a very important property of the method. The algorithm is fast (seconds) and the speed is known in advance, which makes the method suitable for on-line process-monitoring applications. If the noise level of the spectral data is low, GRAM will perform very well. If the noise level is high, GRAM can lead to rough estimates of reaction rate constants [15,16]. In such cases an iterative algorithm will lead to more accurate results. Recently, a new iterative algorithm was developed by Bijlsma *et al* [15,16] in order to obtain accurate estimates of reaction rate constants from noisy data. The algorithm, which is called LM-PAR, consists of a combination of the Levenberg–Marquardt algorithm [34] and alternating least squares steps of the PARAFAC model [35,36], using the results from GRAM as a very good set of starting values for the parameters to be estimated. In this paper, LM-PAR will be termed GRAM-LM-PAR, indicating that the GRAM results are used as starting values for LM-PAR.

Two-way techniques and three-way methods were used to estimate reaction rate constants from chemical reactions in previous papers [11,13,15,16]. In this paper the accuracy and speed of two important curve resolution algorithms, the three-way methods GRAM and GRAM-LM-PAR and two new three-way methods are compared using two different datasets. The first two-way technique is a curve resolution algorithm which uses spectral information that is known in advance. For example, it is possible to use the pure spectra of reacting absorbing species if these are known. The second two-way technique is a curve resolution algorithm which does not use spectral information that is known *a priori*. The curve resolution methods proposed in this paper are not traditional curve resolution methods, because model information, in this case specific kinetic information, is used explicitly. Quality assessment is performed using a jackknife method [37].

As mentioned earlier, two new algorithms are introduced in this paper. GRAM uses two slave datasets created from a master dataset by means of a time shift. Of course, it is also possible to create three or more slave datasets from the master dataset such that there is a special relation between the slave datasets. Using more slave datasets can be advantageous because of noise reduction. Next, two or more slave datasets are stacked. Trilinear linear decomposition (TLD) [38] can be used to decompose the trilinear structure. Hence the difference between the data treatment of GRAM and TLD is that TLD uses more than two slave datasets. TLD is also non-iterative. If the TLD results are available, these can be used as a very good set of starting values for LM-PAR, resulting in a new algorithm, TLD-LM-PAR.

Two different datasets are used to test the performance of the two-way and three-way methods. Dataset 1 consists of SW-NIR spectra of the epoxidation of 2,5-di-*tert*-butyl-1,4-benzoquinone [8,11,39]. Dataset 2 consists of UV-VIS spectra of the consecutive reaction of 3-chlorophenylhydrazonopropane dinitrile with 2-mercaptoethanol [12,13,16,40]. Dataset 1 has a moderate noise level and dataset 2 has a low noise level. Hence both datasets represent extremes of what can be expected in chemical practice.

## 2. THEORY

### 2.1. Notation

Boldface capital characters denote matrices, boldface lower-case characters denote vectors and boldface underlined capital characters denote three-way arrays. Define a matrix  $\mathbf{X}$  with dimensions  $M \times N$ . Then  $\mathbf{X}^T$  represents the transpose of  $\mathbf{X}$ ,  $\mathbf{X}^{-1}$  represents the inverse of  $\mathbf{X}$ ,  $\hat{\mathbf{X}}$  denotes the estimate of  $\mathbf{X}$ ,  $\bar{\mathbf{X}}$  represents the truncated matrix  $\mathbf{X}$  (see later) and  $\|\mathbf{X}\|$  indicates the Frobenius or Euclidean norm of  $\mathbf{X}$ .  $\mathbf{X}^{(j)}$  is the matrix  $\mathbf{X}$  after the  $j$ th iteration,  $\mathbf{X}_i$  is the  $i$ th column of  $\mathbf{X}$ ,  $\mathbf{X}_i^{(j)}$  is the  $i$ th column of  $\mathbf{X}$  after the  $j$ th iteration and  $\mathbf{1}$  represents a vector of ones. The most important properties of the two-way and three-way methods which will be discussed in the following subsections are summarized in Table I.

### 2.2. The model

Let matrix  $\mathbf{X}$  ( $M \times N$ ) be a collection of spectra taken in time of a certain chemical reaction with  $M$  equidistant time points at  $N$  wavelengths,  $K$  reacting absorbing species and uniform errors assumed in the data. Matrix  $\mathbf{X}$  can be expressed as the following equation, assuming the Beer–Lambert law [41]:

$$\mathbf{X} = \mathbf{F}\mathbf{D}^T + \mathbf{E} \quad (1)$$

where  $\mathbf{F}$  ( $M \times K$ ) contains the concentration profiles of  $K$  reacting absorbing species,  $\mathbf{D}$  ( $N \times K$ ) contains the pure spectra of  $K$  reacting absorbing species and  $\mathbf{E}$  ( $M \times N$ ) is a matrix of errors (model errors, experimental errors and instrumental noise).

Consider the two-step first-order consecutive chemical reaction



Table I. Important properties of some two-way and three-way methods: use of *a priori* spectral information; time shift; number of slabs created from master dataset; iterative nature; speed; and applicability for on-line use. In the case of speed, ‘++’ means less than 10 s (fast), ‘+’ means more than 10 s (medium) and ‘–’ means minutes to hours (slow)

Algorithm	<i>A priori</i> information	Time shift	Number of slabs	Iterative	Speed	On-line
CCR	Yes	No	1	Yes	+	No
WCR	No	No	1	Yes	+	No
GRAM	No	Yes	2	No	++	Yes
GRAM-LM-PAR	No	Yes	2	Yes	–	No
TLD	No	Yes	$\geq 3$	No	++	Yes
TLD-LM-PAR	No	Yes	$\geq 3$	Yes	–	No

where  $k_1$  and  $k_2$  are reaction rate constants, both in  $\text{min}^{-1}$ . Equations (2)–(4) are the kinetic rate equations describing the concentration profiles of species U, V and W respectively, assuming that only species U is initially present:

$$C_{U,i} = C_{U,0}e^{-k_1t_i} \quad (2)$$

$$C_{V,i} = \frac{k_1 C_{U,0}}{k_2 - k_1} (e^{-k_1t_i} - e^{-k_2t_i}) \quad (3)$$

$$C_{W,i} = C_{U,0} - C_{U,i} - C_{V,i} \quad (4)$$

where  $C_{U,i}$ ,  $C_{V,i}$  and  $C_{W,i}$  are the concentrations of species U, V and W at time  $t_i$  respectively and  $C_{U,0}$  is the initial concentration of species U at time 0. Equations (2)–(4) are the columns of  $\mathbf{F}$ . Hence the first column of  $\mathbf{F}$  represents Equation (2) for example.

Suppose only  $\mathbf{X}$  is *known* and  $\mathbf{F}$  and  $\mathbf{D}$  are both *unknown*. In this case it is impossible to estimate the reaction rate constants of the considered chemical reaction using techniques which are based on fitting the kinetic expressions to the concentration versus time data, since these concentrations are not measured. Sometimes a part of  $\mathbf{D}$  is *known* in advance. If this is not the case, it is very often possible to record pure spectra of reactants and products in practice. However, obtaining pure spectra of intermediate species can be a problem because of their fast decomposition. Matrix  $\mathbf{F}$  is *unknown*, but a model for  $\mathbf{F}$  (structure) is *known* if a suitable kinetic model for the chemical reaction of interest and the initial concentrations of the different reacting absorbing species are *known*. Iterative algorithms are necessary in order to estimate the reaction rate constants of interest.

### 2.3. Two-way methods

2.3.1. *Classical curve resolution (CCR)*. In this subsection a classical curve resolution approach which uses spectral information that is known in advance will be discussed very briefly. The algorithm has been described extensively in an earlier paper [13] and will be termed classical curve resolution (CCR) in this paper. Define

$$\mathbf{X}^T = \mathbf{D}\mathbf{F}^T + \mathbf{E}^T \quad (5)$$

The goal of CCR is to minimize Equation (6) over  $k_1$  and  $k_2$ , ensuring that for the proper values of the parameters the minimum is achieved:

$$\min_{k_1, k_2, \mathbf{d}_2} \|\mathbf{X}^T - \mathbf{D}\mathbf{F}^T\|^2 \quad (6)$$

where  $\mathbf{D} = [\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3]$  and  $\mathbf{d}_1$  represents the first column of  $\mathbf{D}$  which contains the pure spectrum of the reactant. The vectors  $\mathbf{d}_2$  and  $\mathbf{d}_3$  represent the pure spectra of intermediate and product respectively. The pure spectra of reactant and product ( $\mathbf{d}_1$  and  $\mathbf{d}_3$ ) are both assumed to be known in advance. The reaction rate constants  $k_1$  and  $k_2$  are optimized according to the following steps.

1. Use a starting value for  $k_1$  and  $k_2$ .
2. Reconstruct  $\mathbf{F}$ . This gives  $\hat{\mathbf{F}}$ .
3. Calculate  $\tilde{\mathbf{X}}^T$  according to

$$\tilde{\mathbf{X}}^T = \mathbf{X}^T - \mathbf{d}_1\mathbf{f}_1^T - \mathbf{d}_3\mathbf{f}_3^T \quad (7)$$

4. Estimate  $\hat{\mathbf{d}}_2$  using  $\tilde{\mathbf{X}}^T$ ,  $\hat{\mathbf{f}}_2$  and non-negativity constraints [42].
5. Modify  $k_1$  and  $k_2$  according to the Levenberg–Marquardt algorithm [34] by minimizing Equation (6).
6. Update  $\hat{\mathbf{f}}_2$  and  $\hat{\mathbf{d}}_2$ .
7. Repeat steps 5 and 6 until convergence has been reached.

The two-way algorithm CCR and the least squares parts of the three-way algorithms (see Sections 2.4.3–2.4.5) can account for non-uniform errors present in the data by using weighted least squares (WLS) if the variance of the measurement error is known [43]. However, the use of WLS is beyond the scope of this paper, since the structure of the measurement error is unknown.

**2.3.2. Weighted curve resolution (WCR).** A modified curve resolution procedure which is termed weighted curve resolution (WCR) is described in this subsection. An extensive explanation of the algorithm can be found in an earlier paper [11]. Equation (8) shows the singular value decomposition (SVD) of  $\mathbf{X}^T$  assuming  $M \leq N$ ; for cases where  $N < M$  is valid,  $\mathbf{X}$  can be used instead of  $\mathbf{X}^T$ :

$$\mathbf{X}^T = \mathbf{U}\mathbf{S}\mathbf{V}^T \quad (8)$$

with  $\mathbf{U}^T\mathbf{U} = \mathbf{I}$ ,  $\mathbf{V}^T\mathbf{V} = \mathbf{V}\mathbf{V}^T = \mathbf{I}$ ,  $\mathbf{U}$  ( $N \times M$ ),  $\mathbf{V}$  ( $M \times M$ ) and  $\mathbf{S}$  ( $M \times M$ ) a diagonal matrix with the singular values arranged in decreasing order on the diagonal. Equation (9) shows the truncated version of Equation (8) to the first  $L$  significant singular values:

$$\bar{\mathbf{X}}^T = \bar{\mathbf{U}}_L \bar{\mathbf{S}}_L \bar{\mathbf{V}}_L^T \quad (9)$$

If Equation (5) holds, then  $\hat{\mathbf{F}}$  and  $\bar{\mathbf{V}}_L$  span the same space. Following the target factor analysis (TFA) approach used by Maeder and co-workers [9,18,20], the objective function given in Equation (10) is minimized over  $k_1$  and  $k_2$ , ensuring that for the proper  $k_1$  and  $k_2$  the minimum will be attained:

$$\min_{k_1, k_2} \|\mathbf{I} - \hat{\mathbf{F}}(\hat{\mathbf{F}}^T \hat{\mathbf{F}})^{-1} \hat{\mathbf{F}}^T \bar{\mathbf{V}}_L \bar{\mathbf{S}}_L\|^2 \quad (10)$$

where  $\bar{\mathbf{V}}_L$  and  $\bar{\mathbf{S}}_L$  are both fixed. Because the columns of  $\bar{\mathbf{V}}_L$  are weighted by  $\bar{\mathbf{S}}_L$ , the algorithm is called weighted curve resolution. In this way, account is taken of non-uniform errors present in the data.

WCR does not use any spectral information that is known in advance. Also, the knowledge that  $\hat{\mathbf{D}}$  is non-negative is not used in WCR.

## 2.4. Three-way methods

**2.4.1. The three-way structure.** The three-way structure which is used by the three-way algorithms in order to estimate unknown reaction rate constants is discussed briefly in this subsection. An extensive explanation of the three-way structure used can be found in earlier papers [15,16]. Consider matrix  $\mathbf{X}$  ( $M \times N$ ), with spectra recorded in time of the reacting system mentioned earlier, again. First, a column with constants, e.g. a column with ones,  $(1 \dots 1)^T$ , is added to  $\mathbf{X}$  to construct the augmented matrix  $\mathbf{X}^*$  ( $M \times (N+1)$ ). Reasons for this step can be found in earlier papers [15,16] and are also treated in detail by Windig and Antalek [30]. Next,  $\mathbf{X}^*$  is split into two data matrices  $\mathbf{X}_1^*$  and  $\mathbf{X}_2^*$  using a time shift  $S$ . The matrices  $\mathbf{X}_1^*$  ( $(M-S) \times (N+1)$ )

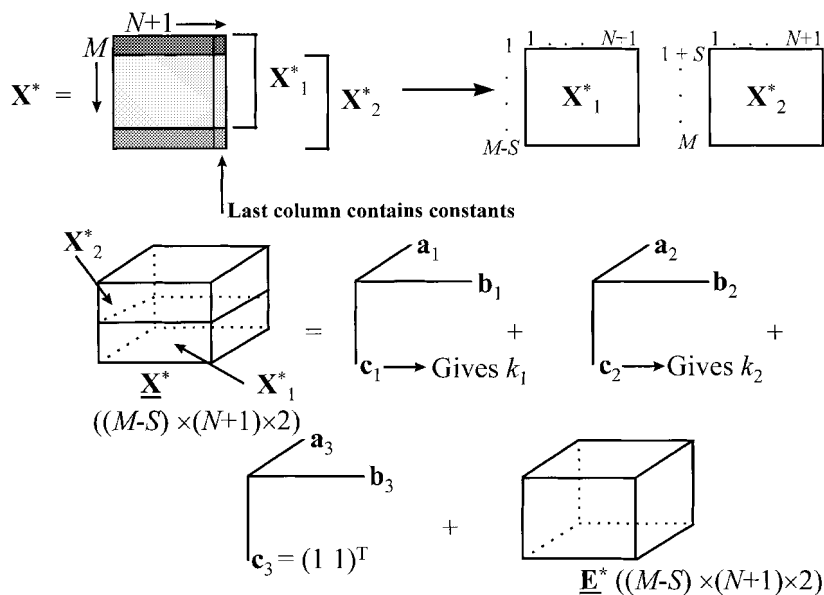


Figure 1. PARAFAC model.

and  $\mathbf{X}_2^*((M-S) \times (N+1))$  which are created by separating  $\mathbf{X}^*$  are used to construct the three-way array  $\underline{\mathbf{X}}^*$   $((M-S) \times (N+1) \times 2)$  by means of stacking and can be modelled with PARAFAC. This is shown in Figure 1.

From  $\underline{\mathbf{X}}^*$  the following three loading matrices can be constructed:

- $\mathbf{A} = [\mathbf{a}_1 \ \mathbf{a}_2 \ \mathbf{a}_3]$   $((M-S) \times 3)$  with  $k$ -rank [44] equal to 3;
- $\mathbf{B} = [\mathbf{b}_1 \ \mathbf{b}_2 \ \mathbf{b}_3]$   $((N+1) \times 3)$  with  $k$ -rank equal to 3;
- $\mathbf{C} = [\mathbf{c}_1 \ \mathbf{c}_2 \ \mathbf{c}_3]$   $(2 \times 3)$  with  $k$ -rank equal to 2, assuming  $k_1 \neq k_2$ .

From the third loading matrix  $\mathbf{C}$  the reaction rate constants  $k_1$  and  $k_2$  can be extracted very easily if the time shift parameter  $S$  is known [15,16]. The structure of  $\mathbf{C}$  is

$$\mathbf{C} = \begin{pmatrix} c_{11} & c_{12} & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad (11)$$

where  $k_1$  and  $k_2$  are  $\ln(c_{11})/S$  and  $\ln(c_{12})/S$  respectively.

So far, the whole procedure is valid if  $\mathbf{X}^*$  is split into two datasets. It is also possible to split  $\mathbf{X}^*$  into more than two datasets using a certain step size. Creating more datasets from  $\mathbf{X}^*$  can be an advantage with respect to noise reduction. It is important to stress that the same master dataset is used over and over again. Assume for simplicity that  $\mathbf{X}^*$  has to be split into three data matrices or slabs  $\mathbf{X}_1^*$ ,  $\mathbf{X}_2^*$  and  $\mathbf{X}_3^*$  using a step size  $G$ . This is visualized in Figure 2. From this figure it can be observed that the step size between  $\mathbf{X}_1^*$  and  $\mathbf{X}_2^*$  is equal to the step size between  $\mathbf{X}_2^*$  and  $\mathbf{X}_3^*$ . The three matrices which are created by separating  $\mathbf{X}^*$  are used to construct the three-way array  $\underline{\mathbf{X}}^*$   $([M - (G \times 3)] \times (N+1) \times 3)$  by means of stacking and can be modelled with PARAFAC similarly to the procedure shown in Figure 1. The loading matrices can be constructed and, from the third loading matrix according to Equation (12), estimates for  $k_1$  and  $k_2$  can be extracted which can be averaged:

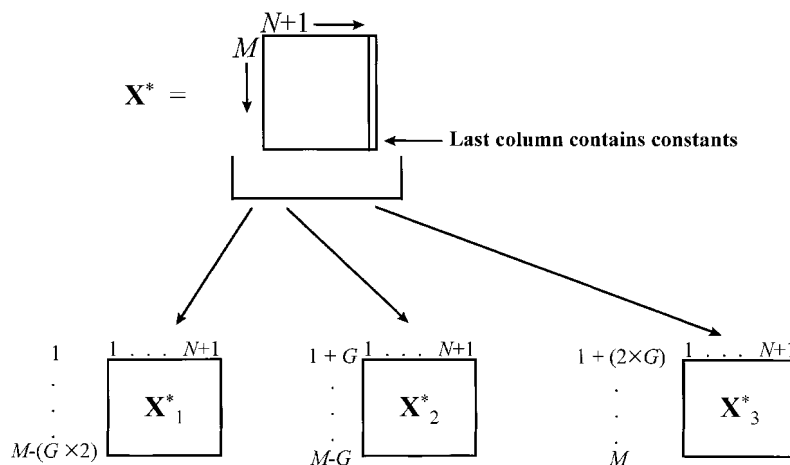


Figure 2. Splitting  $\mathbf{X}^*$  into three data matrices  $\mathbf{X}_1^*$ ,  $\mathbf{X}_2^*$  and  $\mathbf{X}_3^*$ .

$$\mathbf{C} = \begin{pmatrix} c_{11} & c_{12} & 1 \\ c_{21} & c_{22} & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad (12)$$

where  $\ln(c_{11})/\ln(c_{21}) = \ln(c_{21}) = k_1$  and  $\ln(c_{12})/\ln(c_{22}) = \ln(c_{22}) = k_2$  for the noiseless case. Hence two estimates for both  $k_1$  and  $k_2$  are obtained if three slabs are created from  $\mathbf{X}^*$ . These estimates can be averaged. Likewise, three estimates of reaction rate constants are obtained from four slabs, etc.

2.4.2. *The generalized rank annihilation method (GRAM).* Details about the GRAM algorithm used in this paper can be found in earlier papers [15,16]. Here only the essential steps are given. Two slabs are created from  $\mathbf{X}^*$  by means of a time shift  $S$ .

1. Use a generalized eigenvalue problem (GEP) to decompose the three-way array  $\underline{\mathbf{X}}^*$  ( $(M - S) \times (N + 1) \times 2$ ) into three loading matrices  $\mathbf{A}^{(0)}$ ,  $\mathbf{B}^{(0)}$  and  $\mathbf{C}^{(0)}$ . The GEP was solved for  $\mathbf{X}_1^*$  and  $\mathbf{X}_2^*$  according to the procedure described by Wilson *et al* [33].
2. Estimate the reaction rate constants from the scaling factors listed in  $\mathbf{C}^{(0)}$  if the time shift  $S$  is known.

2.4.3. *The Levenberg–Marquardt algorithm and alternating least squares steps of the PARAFAC model (GRAM-LM-PAR).* Details about the algorithm can be found in earlier papers [15,16]. The algorithm will be termed GRAM-LM-PAR instead of LM-PAR, in order to indicate that GRAM results are used as an initial set of starting values.

1. Apply GRAM. This gives  $\mathbf{A}^{(0)}$ ,  $\mathbf{B}^{(0)}$ ,  $\mathbf{C}^{(0)}$ ,  $\hat{k}_1$  and  $\hat{k}_2$ .
2. Define  $\tilde{\mathbf{A}}^{(0)}$ ,  $\tilde{\mathbf{B}}^{(0)}$  and  $\tilde{\mathbf{C}}^{(0)}$ . The first column of  $\mathbf{A}^{(0)}$  represents  $he^{-k_1 t}$  and the second column of  $\mathbf{A}^{(0)}$  represents  $he^{-k_2 t}$  [15,16]. The parameter  $h$  is a constant. Hence  $k_1$  and  $k_2$  can be replaced by  $\hat{k}_1$  and  $\hat{k}_2$  respectively, resulting in  $\tilde{\mathbf{A}}^{(0)}$  which is a reconstruction of the concentration profiles using the reaction rate constant estimates from  $\mathbf{C}^{(0)}$ .  $\tilde{\mathbf{C}}^{(0)}$  is a reconstruction of  $\mathbf{C}^{(0)}$  where the

last row and last column are set to a row vector and column vector of ones respectively. This is known in advance and hence this information has to be used.  $\tilde{\mathbf{B}}^{(0)}$  is constructed by means of a three-way least squares fit (PARAFAC fit) from  $\mathbf{X}^*$ ,  $\tilde{\mathbf{C}}^{(0)}$  and  $\tilde{\mathbf{A}}^{(0)}$ .

3. Subtract the contribution of the column of constants from  $\tilde{\mathbf{X}}^*$ , resulting in  $\dot{\mathbf{X}}^*$ , where  $\tilde{\mathbf{X}}^*$  is created by unfolding  $\mathbf{X}^*$ :

$$\dot{\mathbf{X}}^{*(0)} = \tilde{\mathbf{X}}^* - \tilde{\mathbf{a}}_3^{(0)}(\tilde{\mathbf{c}}_3^{(0)\text{T}} | \otimes | \tilde{\mathbf{b}}_3^{(0)\text{T}}) \quad (13)$$

where  $| \otimes |$  denotes the Khatri–Rao product [45].

4. Next, matrices are partitioned:

$$\tilde{\mathbf{A}}^{(0)} = [\tilde{\mathbf{A}}^{(0)}, \mathbf{a}_3^{(0)}] \quad (14)$$

$$\tilde{\mathbf{B}}^{(0)} = [\tilde{\mathbf{B}}^{(0)}, \tilde{\mathbf{b}}_3^{(0)}] \quad (15)$$

$$\tilde{\mathbf{C}}^{(0)} = \left[ \tilde{\mathbf{C}}^{(0)}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right] \quad (16)$$

where the matrices with a double tilde consist of the first two columns of the considered matrix.

5. Equation (17) is minimized using the Levenberg–Marquardt algorithm [34]:

$$\min_{k_1, k_2} \|\dot{\mathbf{X}}^{*(0)} - \tilde{\mathbf{A}}^{(0)}(\tilde{\mathbf{C}}^{(0)\text{T}} | \otimes | \tilde{\mathbf{B}}^{(0)\text{T}})\|^2 \quad (17)$$

Update  $\tilde{\mathbf{A}}^{(0)}$  and  $\tilde{\mathbf{C}}^{(0)}$  using the optimal values for  $k_1$  and  $k_2$  according to the Levenberg–Marquardt algorithm.

6. Update  $\tilde{\mathbf{B}}^{(0)}$  using a least squares PARAFAC step.
7. Repeat steps 3–6 until the relative change in fit between two iterations is below  $10^{-6}$ .

It is important to state that GRAM-LM-PAR can only be used if the exponential functions describing the concentration profiles of the different species can be written as a sum of exponential functions [15,16]. This also holds for GRAM. Hence, in kinetics, GRAM and GRAM-LM-PAR can only be used in order to model (pseudo-)first-order kinetics.

**2.4.4. Trilinear decomposition (TLD) algorithm.** The trilinear decomposition (TLD) algorithm is well described in the literature by Booksh *et al* [38]. Here a short description is given.

1. Assume that three slabs are created from  $\mathbf{X}^*$ . Decompose the three-way array  $\mathbf{X}^*$  ( $[M - (G \times 3)] \times (N + 1) \times 3$ ) into three loading matrices  $\mathbf{A}^{(0)}$ ,  $\mathbf{B}^{(0)}$  and  $\mathbf{C}^{(0)}$ . A common space of  $\mathbf{X}_1^* + \mathbf{X}_3^*$  was used in the case of three slabs and a common space of  $\mathbf{X}_1^* + \mathbf{X}_4^*$  was used in the case of four slabs.
2. Estimate the reaction rate constants from the scaling factors listed in  $\mathbf{C}^{(0)}$ .
3. Average the estimates for  $k_1$  and  $k_2$ . This will yield mean estimated  $k_1$  and  $k_2$ .

**2.4.5. The Levenberg–Marquardt algorithm and alternating least squares steps of the PARAFAC model (TLD-LM-PAR).** The TLD-LM-PAR algorithm is exactly the same algorithm as GRAM-LM-PAR, except that in TLD-LM-PAR the TLD results are used as an initial set of starting values for  $k_1$  and  $k_2$  instead of the GRAM results for the reaction rate constants. Only the first two steps



of the TLD-LM-PAR algorithm differ from those of the GRAM-LM-PAR algorithm. For TLD-LM-PAR the first two steps are given below.

1. Create a three-way structure from a desired number of matrices formed from  $\underline{\mathbf{X}}^*$ . Next, apply TLD. This gives  $\mathbf{A}^{(0)}$ ,  $\mathbf{B}^{(0)}$ ,  $\mathbf{C}^{(0)}$  and multiple reaction rate constant estimates. The reaction rate constant estimates are averaged, resulting in  $\hat{k}_1$  and  $\hat{k}_2$ .
2. Define  $\tilde{\mathbf{A}}^{(0)}$ ,  $\tilde{\mathbf{B}}^{(0)}$  and  $\tilde{\mathbf{C}}^{(0)}$ . The first column of  $\mathbf{A}^{(0)}$  represents  $he^{-k_1t}$  and the second column of  $\mathbf{A}^{(0)}$  represents  $he^{-k_2t}$  [15,16]. Hence  $k_1$  and  $k_2$  can be replaced by  $\hat{k}_1$  and  $\hat{k}_2$  respectively, resulting in  $\tilde{\mathbf{A}}^{(0)}$  which is a reconstruction of the concentration profiles using the averaged reaction rate constant estimates.  $\tilde{\mathbf{C}}^{(0)}$  is a reconstruction of  $\mathbf{C}^{(0)}$  where the last row and last column are set to a row vector and column vector of ones respectively, and  $\hat{k}_1$  and  $\hat{k}_2$  are used for the reconstruction of  $\mathbf{C}^{(0)}$ .  $\tilde{\mathbf{B}}^{(0)}$  is constructed by means of a three-way least squares fit (PARAFAC fit) from  $\underline{\mathbf{X}}^*$ ,  $\tilde{\mathbf{C}}^{(0)}$  and  $\tilde{\mathbf{A}}^{(0)}$ .

The next steps are equal to the GRAM-LM-PAR algorithm.

### 2.5. Quality assessment of the estimated reaction rate constants

In previous work a procedure for quality assessment of the estimated reaction rate constants was discussed intensively [11,13,15,16]. Here some important aspects are explained. Assume that  $n$  individual batch process runs are performed. Hence  $n$  individual estimates of reaction rate constants can be obtained. Fluctuations between individual estimates of reaction rate constants are caused by model errors, which are assumed to be small if the kinetic model used and the Beer–Lambert law are valid, experimental errors and instrumental noise. If a standard deviation is calculated of the reaction rate constant estimates, this can be represented as the upper error bound of estimates of the reaction rate constants, when variations due to experimental errors and instrumental noise are considered. Experimental noise can be caused by concentration errors and errors due to the start of the reaction, for example. Instrumental noise is caused by variations of the instrument.

A lower error bound can also be defined which mainly represents the consequences of instrumental noise. A jackknife procedure [37] can be used in order to estimate the lower error bound [11,13,15,16]. Assume that 21 spectra are recorded in time of a certain chemical process at  $N$  wavelengths and assume a jackknife interval of four spectra. Firstly, spectra 1, 5, 9, 13 and 17 are left out and the remaining spectra are used to estimate the reaction rate constants. Secondly, spectra 1, 5, 9, 13 and 17 are left in and spectra 2, 6, 10, 14 and 18 are left out and the remaining spectra are used again to estimate the reaction rate constants. Finally, this will lead to a certain number of estimates of reaction rate constants. The jackknife standard deviation obtained represents the lower error bound of mainly instrumental noise.

### 2.6. Reproducibility

The reproducibility for  $n$  batch process runs was obtained using the equation

$$\text{reproducibility} = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n \|\mathbf{X}_i - \mathbf{X}_m\|^2}}{\|\mathbf{X}_m\|} \times 100\% \quad (18)$$

where  $\mathbf{X}_i$  is the spectral matrix for batch process run  $i$  and  $\mathbf{X}_m$  is the averaged spectral matrix obtained from  $n$  individual batch process runs.

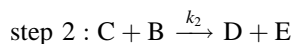
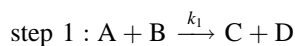
### 2.7. Correlation coefficients

The correlation coefficient  $r$  between the  $k_1$  and  $k_2$  estimates was obtained for each algorithm. The value for  $r$  gives an indication about the quality of the algorithm used. A small value of  $r$  indicates that the algorithm can distinguish between  $k_1$  and  $k_2$  during the optimization procedure very easily. Hence the reaction rate constant estimates are more reliable. The value of  $r$  can be positive or negative. If  $r$  is negative, this means that  $k_1$  and  $k_2$  are negatively correlated. A larger value of  $k_1$  is associated with a smaller value of  $k_2$  and *vice versa*.

## 3. EXPERIMENTAL

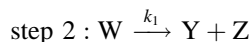
### 3.1. The reaction

**3.1.1. Dataset 1.** Dataset 1 consisted of short-wavelength near-infrared (SW-NIR) spectra taken in time of the epoxidation of 2,5-di-*tert*-butyl-1,4-benzoquinone using *tert*-butyl hydroperoxide and Triton B catalyst [8,11,39]. The reaction consisted of two steps:



Species A, B, C, D and E are specified in Table II. No distinction was made between the *cis* and *trans* product ( $E$ ) of the second step of the reaction, because the spectral differences are negligible in SW-NIR. If species B is present in large excess, the first and second steps of the reaction both become pseudo-first-order instead of second-order. Hence, Equations (2)–(4) can be used in order to describe the concentration profiles of the reactant (A), intermediate (C) and main product (E) of the reaction respectively. Species A, C and E were monitored.

**3.1.2. Dataset 2.** Dataset 2 consisted of UV-VIS recorded spectra of the consecutive reaction of 3-chlorophenylhydrazonopropane dinitrile (U) with 2-mercaptoethanol (V) [12,13,16,40]. The reaction consisted of two steps:



Species U, V, W, Y and Z are specified in Table II. If V is present in large excess, the first step

Table II. Species involved in dataset 1 and dataset 2

Dataset 1		Dataset 2	
A	2,5-di- <i>tert</i> -butyl-1,4-benzoquinone	U	3-Chlorophenylhydrazonopropane dinitrile
B	<i>Tert</i> -butyl hydroperoxide	V	2-Mercaptoethanol
C	2,5-di- <i>tert</i> -butyl mono-epoxide-1,4-benzoquinone	W	Intermediate adduct
D	<i>Tert</i> -butyl alcohol	Y	3-Chlorophenylhydrazonocynoacetamide
E	<i>Cis</i> and <i>trans</i> 2,5-di- <i>tert</i> -butyl di-epoxide-1,4-benzoquinone	Z	Ethylene sulphide

of the reaction becomes pseudo-first-order instead of second-order. Hence, Equations (2)–(4) can be used in order to describe the concentration profiles of U, W and Y respectively. In this paper, species U (reactant), W (intermediate) and Y (product) were monitored.

### 3.2. Sample preparation

**3.2.1. Dataset 1.** First, the cuvette was filled with 0.264 g (1.2 mmol) of 2,5-di-*tert*-butyl-1,4-benzoquinone, as synthesized by the procedure described by Hairfield *et al* [39], which was then dissolved in 15 ml dioxane (Acros 99+%). Next, 1.55 ml (12 mmol) of a *tert*-butylhydroperoxide 70% (Acros) solution in water and 13.21 ml ethanol (BDH Laboratory Supplies, pro analysis) were both added. After the target reaction temperature had been reached, data collection was started immediately after addition of 0.24 ml ice-cold Triton B catalyst (Acros, 40 wt% in methanol) in 0.50 ml ethanol and 0.60 ml dioxane. The excess of *tert*-butylhydroperoxide in moles was 10 times 2,5-di-*tert*-butyl-1,4-benzoquinone. The experiment was repeated eight times. More details about the sample preparation can be found in an earlier paper [11].

**3.2.2. Dataset 2.** The cuvette was filled with 2.5 ml of a working solution containing 51.71  $\mu\text{mol l}^{-1}$  3-chlorophenylhydrazonopropane dinitrile (Acros, 99+%) buffered with  $\text{KH}_2\text{PO}_4$  (Acros, pro analysis 0.2 mol  $\text{l}^{-1}$ , pH 4.4). When the target reaction temperature had been reached, data collection was started by addition of 10  $\mu\text{l}$  of a solution of 2-mercaptoethanol (35.65  $\mu\text{mol}$ ). This solution consisted of 250  $\mu\text{l}$  pure 2-mercaptoethanol (Acros 99%) and 750 ml  $\text{KH}_2\text{PO}_4$  buffer solution. Hence the excess of 2-mercaptoethanol in moles was 276 times 3-chlorophenylhydrazonopropane dinitrile. The experiment was repeated 10 times. More details about the sample preparation can be found in an earlier paper [16].

### 3.3. Experimental set-up

For both datasets the same experimental set-up was used. A Hewlett Packard 8453 UV-VIS spectrophotometer with diode array detection was used to measure spectra of the reacting system. For dataset 1 a quartz cuvette with 10.00 cm path length (Hellma Benelux) was used to obtain spectra of the reaction mixture. For dataset 2 a quartz cuvette with 1.00 cm path length (Hellma Benelux) was used. A Pt-100 thermocouple and a constant-temperature bath (Neslab) were used to control the temperature inside the cuvette. The experimental conditions for dataset 1 and dataset 2 were different and are summarized in Table III. Details about the experimental set-up can be found in earlier papers [11,16].

Table III. Experimental conditions used for dataset 1 and dataset 2

Experimental condition	Dataset 1	Dataset 2
Reaction temperature ( $^{\circ}\text{C}$ )	17	25
Integration time (s)	1	1
Sampling time (s)	5	10
Total run time (s)	1200	2700
Wavelength range (nm)	800–1100	200–600
Wavelength interval (nm)	1.0	1.0
Number of recorded spectra	241	271

### 3.4. Dataprocessing

3.4.1. *Dataset 1.* Second-derivative spectra were estimated using a Savitzky–Golay filter [46] with a window size of 15 data points. To stress the spectral features of the appearing and disappearing species, second-derivative difference spectra were calculated after subtracting the fourth recorded spectrum from all the other spectra remaining. Hence the first three recorded spectra were not used for data processing because of the moderate reproducibility of these spectra as described in an earlier paper [11]. The small wavelength range 860–880 nm was used for data processing. If this wavelength range is used, the spectral features are caused by the three species which were monitored (species A, C and E [8,11]). For WCR the data matrix with recorded spectra of the reacting system was truncated to three singular values. More details about data processing can be found in an earlier paper [11]. For WCR, starting values of 0.30 and 0.05  $\text{min}^{-1}$  were used for  $k_1$  and  $k_2$  respectively.

3.4.2. *Dataset 2.* A spectrum of  $\text{KH}_2\text{PO}_4$  buffer solution was used as blank. The wavelength range 300–500 nm was used for data processing. Using this spectral range, there are only spectral features caused by species U, W and Y. For CCR and WCR, starting values of 0.30 and 0.05  $\text{min}^{-1}$  were used for  $k_1$  and  $k_2$  respectively.

### 3.5. Pure spectra

For dataset 2 the pure spectrum of the reactant was recorded by averaging seven recorded spectra just before addition of species V. In order to record the pure spectrum of the product, the reaction mixture was allowed to react for 8 h. After this period, seven spectra of the reaction mixture were taken. The recorded spectra represented the pure spectrum of the product, because the concentrations of reactant and intermediate will be negligible after 8 h. The pure spectrum of the intermediate could not be recorded, since it is not available at chemical suppliers and hard to synthesize in its pure form. For dataset 1, no pure spectra of species were available. Hence it was not possible to apply the CCR algorithm for dataset 1.

## 4. RESULTS AND DISCUSSION

### 4.1. Reproducibility

The reproducibility was calculated for both datasets. For dataset 1 the reproducibility of the second-derivative difference spectra was equal to 23.38%. This moderate reproducibility is mainly a consequence of the small differences in absorbances of the species and the error propagation caused by taking the second derivative. For dataset 2 the reproducibility of the blank-corrected spectra was equal to 0.54%. Hence there is a large difference in reproducibility between dataset 1 and dataset 2. For both datasets the mean batch process run was obtained by averaging the spectra of the individual batch process runs. The spectra of the mean batch process run are plotted in Figure 3 for dataset 1 and dataset 2.

### 4.2. Estimates of reaction rate constants from individual batch processes

4.2.1. *Dataset 1.* For eight individual batch process runs the reaction rate constants were estimated using different algorithms. The results are listed in Table IV. The individual standard deviations represent the upper error bounds. The correlation coefficients give an indication about the performance of the different algorithms used. For GRAM and GRAM-LM-PAR the time shift was equal to 30 spectra. From previous work [15] it appeared that this time shift led to the lowest

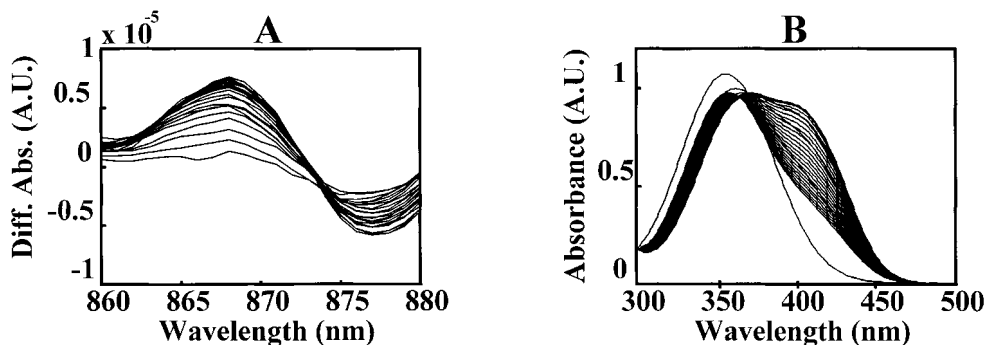


Figure 3. Spectra for mean batch process run: A, dataset 1; B, dataset 2. The spectra are plotted using an interval of 10 spectra.

standard deviation for  $k_1$  and  $k_2$ . For TLD and TLD-LM-PAR the step size was equal to 20 spectra. It appeared that these values led to the lowest individual standard deviations. From Table IV the following aspects can be observed.

1. The lowest individual standard deviation (upper error bound) for both reaction rate constants is obtained using WCR.
2. All algorithms led to approximately the same individual standard deviation for  $k_1$ .
3. The magnitude of the individual standard deviation for  $k_2$  is very dependent on the algorithm used.
4. There is no gain in individual standard deviation for both reaction rate constants if more than two slabs are used. The use of TLD and TLD-LM-PAR led to no improvement in individual standard deviations compared to GRAM and GRAM-LM-PAR respectively. Hence there is no noise reduction if more than two slabs are used. Note that the same data are used over and over again if slabs are created. Hence the same noise realization of the data is used again and again.
5. GRAM-LM-PAR and TLD-LM-PAR led to an improvement in individual standard deviation for  $k_2$  compared to GRAM and TLD respectively. For  $k_1$  the individual standard deviations will be slightly larger. Hence, with respect to  $k_2$ , performing iterations after GRAM or TLD is advantageous. Note that both GRAM and TLD are not least squares methods.

Table IV. Results for dataset 1: mean estimated reaction rate constants over eight  $k_1$  and  $k_2$  estimates for eight repeated individual batch processes obtained with different methods. The standard deviations (STDs) obtained for the individual batch processes are also given. For WCR the starting values for  $k_1$  and  $k_2$  were 0.30 and 0.05  $\text{min}^{-1}$  respectively

Method	Mean $k_1$ ( $\text{min}^{-1}$ )	STD $k_1$ ( $\text{min}^{-1}$ )	Mean $k_2$ ( $\text{min}^{-1}$ )	STD $k_2$ ( $\text{min}^{-1}$ )	$r$
WCR <sup>a</sup>	0.26	0.03	0.07	0.03	-0.61
GRAM (time shift = 30)	0.28	0.03	0.10	0.06	-0.47
GRAM-LM-PAR (time shift = 30)	0.25	0.04	0.09	0.04	-0.79
TLD, 3 slabs (step size = 20)	0.25	0.03	0.10	0.06	-0.86
TLD, 4 slabs (step size = 20)	0.22	0.03	0.11	0.10	-0.93
TLD-LM-PAR, 3 slabs (step size = 20)	0.24	0.05	0.09	0.04	-0.87

<sup>a</sup> If different sets of starting values were used, similar results were obtained.

6. With respect to  $k_1$  the non-iterative GRAM algorithm gives approximately the same individual standard deviation as other algorithms.
7. With respect to the calculated correlation coefficients, GRAM gives the lowest value, indicating that GRAM can distinguish between  $k_1$  and  $k_2$  very well.
8. Because the noise level of the data is high, iterative algorithms will improve the individual standard deviations obtained with non-iterative algorithms. Hence precise estimates of reaction rate constants in the case of data with a moderate or high noise level are only possible using iterative least squares algorithms.

4.2.2. *Dataset 2.* For 10 individual batch processes the reaction rate constants were estimated using different algorithms. The results are listed in Table V. For GRAM and GRAM-LM-PAR a time shift of 30 spectra was used. From previous work [16] it appeared that this time shift led to the lowest standard deviation for  $k_1$  and  $k_2$ . For TLD and TLD-LM-PAR the step size was equal to 15 spectra. It appeared that these values led to the lowest individual standard deviations. The correlation coefficients between the  $k_1$  and  $k_2$  estimates are also given in Table V. From Table V the following aspects can be observed.

1. Using CCR, the lowest individual standard deviation for  $k_2$  is obtained. Hence using spectral information that is known *a priori* is advantageous. The lowest individual standard deviation for  $k_1$  is obtained using TLD and three slabs. However, the differences between the individual standard deviations obtained with different algorithms are very small. A compromise in individual standard deviation (upper error bound) for  $k_1$  and  $k_2$  is obtained with CCR. Using this method, the individual standard deviation for  $k_1$  is acceptable.
2. All algorithms led to approximately the same order of magnitude of the individual standard deviation for  $k_1$ .
3. There is hardly a gain in individual standard deviation for both reaction rate constants if three-way algorithms with more than two slabs are used. The noise reduction effect is very small.
4. GRAM-LM-PAR and TLD-LM-PAR led to no improvement in individual standard deviation for both reaction rate constants compared to GRAM and TLD respectively. This is caused by the very low noise level of the data.
5. With respect to  $k_1$  and  $k_2$  the non-iterative GRAM algorithm gives approximately the same

Table V. Results for dataset 2: mean estimated reaction rate constants over ten  $k_1$  and  $k_2$  estimates for 10 repeated individual batch process runs obtained with different methods. The standard deviations (STDs) obtained for the individual batch process runs are also given. For CCR and WCR the starting values for  $k_1$  and  $k_2$  were 0.30 and 0.05  $\text{min}^{-1}$  respectively

Method	Mean $k_1$ ( $\text{min}^{-1}$ )	STD $k_1$ ( $10^{-2}$ $\text{min}^{-1}$ )	Mean $k_2$ ( $\text{min}^{-1}$ )	STD $k_2$ ( $10^{-2}$ $\text{min}^{-1}$ )	$r$
CCR <sup>a</sup>	0.32	0.98	0.03	0.02	-0.54
WCR <sup>a</sup>	0.31	0.98	0.03	0.11	-0.82
GRAM (time shift = 30)	0.30	0.83	0.03	0.10	-0.71
GRAM-LM-PAR (time shift = 30)	0.31	0.92	0.03	0.11	-0.79
TLD, 3 slabs (step size = 15)	0.30	0.82	0.03	0.10	-0.78
TLD, 4 slabs (step size = 15)	0.30	0.93	0.03	0.09	-0.82
TLD-LM-PAR, 3 slabs (step size = 15)	0.32	0.91	0.03	0.11	-0.78

<sup>a</sup> If different sets of starting values were used, similar results were obtained.

individual standard deviation as other algorithms. Only CCR led to a lower individual standard deviation for  $k_2$ .

6. With respect to the calculated correlation coefficients, CCR gives the lowest value, indicating that CCR can distinguish between  $k_1$  and  $k_2$  very well, because spectral information that is known *a priori* is used by the algorithm.
7. It seems that GRAM and TLD lead to a systematic difference of the mean value for both reaction rate constants. From the literature it is known that GRAM and TLD can lead to biased estimates [47].
8. Because the noise level of the data is low, iterative least squares algorithms will hardly improve the individual standard deviations obtained with non-iterative algorithms. Hence in the case of data with a low noise level it is not necessary to perform tedious iterations.

#### 4.3. Jackknife results (lower error bounds)

4.3.1. *Dataset 1.* In order to obtain lower error bounds, a jackknife procedure as described in Section 2.5 was applied. Using a certain jackknife interval, a fixed number of spectra were left out. With the leftover spectra the reaction rate constants were estimated using different algorithms. Using three-way methods, the number of spectra left will be less compared to two-way methods because of the applied time shift or step size. The number of leftover spectra must be the same for every algorithm for a fair comparison between lower error bounds obtained with different algorithms. This can be achieved by applying different jackknife intervals. Hence each algorithm applied has a specific jackknife interval resulting in the same number of leftover spectra.

The jackknife results for dataset 1 are listed in Table VI. From this table the following aspects can be observed.

1. GRAM-LM-PAR gives a compromise in lower error bound estimate for both reaction rate constants. Using this method, the lowest lower error bound estimate for  $k_2$  is obtained and an acceptable lower error bound estimate for  $k_1$ . The lowest lower error bound estimate for  $k_1$  is obtained with WCR.
2. The use of GRAM-LM-PAR and TLD-LM-PAR will always lead to an improvement in lower error bound compared to the lower error bound estimates obtained with GRAM and TLD

Table VI. Jackknife results for mean batch process for dataset 1 obtained for different methods. The mean reaction rate constants obtained are the means over 10 jackknife estimates. For WCR the starting values for  $k_1$  and  $k_2$  were 0.30 and 0.05  $\text{min}^{-1}$  respectively

Method	Jackknife interval	Mean $k_1$ ( $\text{min}^{-1}$ )	STD $k_1$ ( $10^{-2} \text{min}^{-1}$ )	Mean $k_2$ ( $\text{min}^{-1}$ )	STD $k_2$ ( $10^{-2} \text{min}^{-1}$ )
WCR <sup>a</sup>	31	0.26	0.10	0.07	0.18
GRAM (time shift = 30)	26	0.28	0.20	0.07	0.16
GRAM-LM-PAR (time shift = 30)	26	0.27	0.11	0.07	0.15
TLD, 3 slabs (step size = 20)	24	0.25	0.18	0.08	0.22
TLD-LM-PAR, 3 slabs (step size = 20)	24	0.26	0.08	0.08	0.22

<sup>a</sup> If different sets of starting values were used, similar results were obtained.

Table VII. Jackknife results for mean batch process for dataset 2 obtained for different methods. The mean reaction rate constants obtained are the means over 10 jackknife estimates. For CCR and WCR the starting values for  $k_1$  and  $k_2$  were 0.30 and 0.05  $\text{min}^{-1}$  respectively

Method	Jackknife interval	Mean $k_1$ ( $\text{min}^{-1}$ )	STD $k_1$ ( $10^{-3} \text{min}^{-1}$ )	Mean $k_2$ ( $\text{min}^{-1}$ )	STD $k_2$ ( $10^{-4} \text{min}^{-1}$ )
CCR <sup>a</sup>	27	0.32	0.03	0.03	0.01
WCR <sup>a</sup>	27	0.32	0.12	0.03	0.13
GRAM (time shift = 30)	25	0.30	0.15	0.03	0.42
GRAM-LM-PAR (time shift = 30)	25	0.31	0.10	0.03	0.11
TLD, 3 slabs (step size = 15)	25	0.30	0.24	0.03	0.39
TLD-LM-PAR, 3 slabs (step size = 15)	25	0.31	0.09	0.03	0.10

<sup>a</sup> If different sets of starting values were used, similar results were obtained.

respectively. This is because of the high noise level of the data. The gain in jackknife standard deviation is approximately a factor two for  $k_1$  if GRAM-LM-PAR is used instead of GRAM. For  $k_2$  there is hardly a gain in jackknife standard deviation if GRAM-LM-PAR is used instead of GRAM.

4.3.2. *Dataset 2.* The jackknife results for dataset 2 are listed in Table VII. Because the number of spectra of dataset 2 is different from that of dataset 1, the applied jackknife interval for dataset 2 differs from the jackknife intervals for dataset 1. From Table VII the following aspects can be observed.

1. CCR gives the lowest estimates for the lower error bound.
2. The use of GRAM-LM-PAR and TLD-LM-PAR results in a gain in lower error bound estimates compared to those obtained with GRAM and TLD respectively. The gain in jackknife standard deviation is approximately a factor two for  $k_1$  and a factor of four for  $k_2$  if GRAM-LM-PAR is used instead of GRAM.

#### 4.4. Estimates of pure spectra

For the  $k_1$  and  $k_2$  estimates from Table V for CCR, WCR and GRAM the concentration profiles were reconstructed. This was not done for the  $k_1$  and  $k_2$  estimates obtained using TLD, GRAM-LM-PAR and TLD-LM-PAR, because the reaction rate constants estimated were very similar to those obtained with GRAM, WCR and CCR respectively. From the spectra of the mean batch process run and the reconstructed concentration profiles the pure spectra of reactant, intermediate and product were estimated for dataset 2 by means of a least squares step using non-negativity constraints. The estimated pure spectra for the reactant and product were subtracted from the recorded spectra of those species. The difference spectra are given in Figures 4 and 5 for the reactant and product respectively. From Figures 4 and 5 it can be observed that there is some structure present in the difference spectra, but the residuals are quite small. It was not possible to compare the estimated pure spectrum for the intermediate with the recorded spectra of the intermediate.



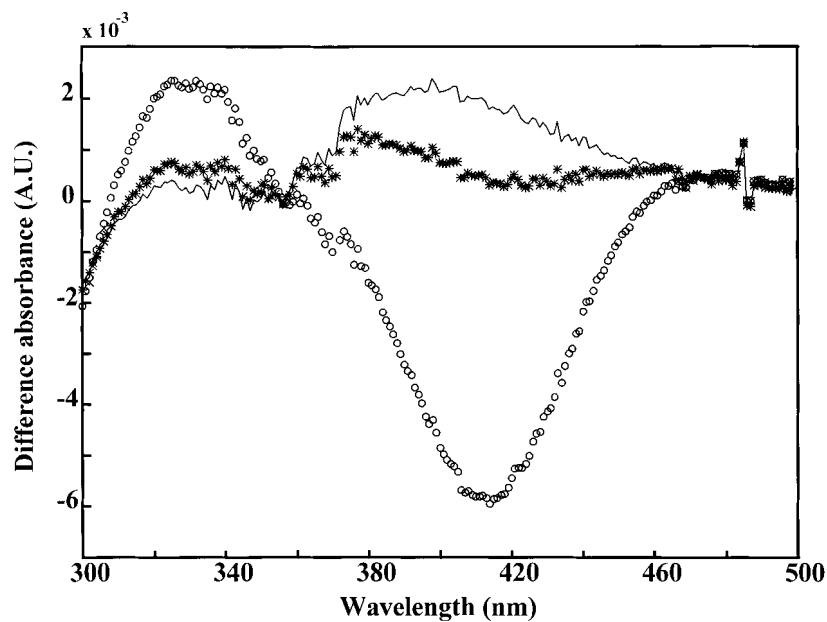


Figure 4. Difference between recorded pure spectrum and estimated pure spectrum for reactant using mean over individual  $k_1$  and  $k_2$  estimates obtained with CCR (—), WCR (\*) and GRAM (○).

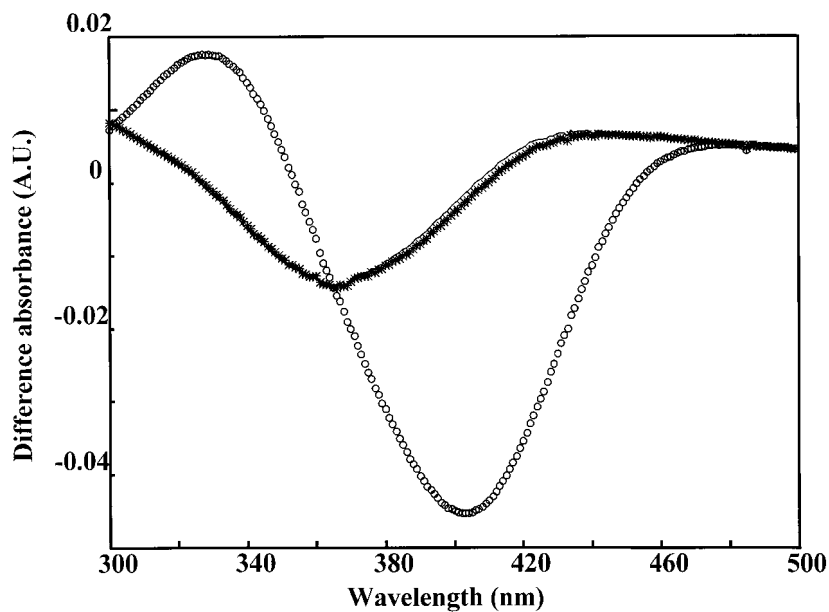


Figure 5. Difference between recorded pure spectrum and estimated pure spectrum for product using mean over individual  $k_1$  and  $k_2$  estimates obtained with CCR (—), WCR (\*) and GRAM (○).

## 5. CONCLUSIONS

In this paper, two-way and three-way methods to estimate reaction rate constants from spectral data of chemical reactions are compared. Two different spectral datasets are used. Dataset 1 consisted of SW-NIR spectra taken in time of a two-step reaction. Dataset 1 showed moderate reproducibility. Dataset 2 consisted of UV-VIS spectra taken in time of another two-step reaction. Dataset 2 showed excellent reproducibility.

The iterative two-way method of weighted curve resolution (WCR) led to the lowest individual standard deviations for the reaction rate constants for dataset 1. The generalized rank annihilation method (GRAM) performed very well, but the individual standard deviations are reduced if iterations are performed after the GRAM solution. This is because of the moderate noise level of dataset 1. The two-way method classical curve resolution (CCR) led to the best values for the individual standard deviations for dataset 2. Using CCR, it is possible to incorporate spectral information that is known beforehand into the algorithm. Also, non-negativity constraints with respect to the pure spectra of the reacting absorbing species involved can be incorporated. The WCR algorithm does not use this kind of spectral information. GRAM led to approximately the same individual standard deviations for the reaction rate constants as iterative methods. This is because of the low noise level of dataset 2.

In general, two-way methods will lead to lower individual standard deviations of reaction rate constant estimates compared to three-way methods. For every specific dataset a certain two-way algorithm will perform the best. The two-way methods CCR and WCR are both of medium speed. It takes more than 10 s to optimize the reaction rate constants if starting values are used which are close to the 'true' values. Because of the iterative nature of both WCR and CCR, the exact speed of the algorithm needed to reach the optimal set of parameters is not known beforehand. This makes CCR and WCR less suitable for on-line estimation of reaction rate constants.

In practice it is very often not necessary to know the values of reaction rate constants very precisely. Hence it is not always necessary to use iterative algorithms. GRAM can be very convenient to use in that case, because the speed of this algorithm is known in advance. This makes GRAM extraordinarily suitable for on-line monitoring of reaction rate constants. Both datasets in this paper showed good results using GRAM and hence this method seems to be the best choice for on-line (pseudo-)first-order kinetic problems. For noisy datasets, more reliable estimates of reaction rate constants can be obtained off-line using iterative algorithms.

## REFERENCES

1. Brown SD. Chemometrics. *Anal. Chem.* 1990; **62**: 84R–101R.
2. Otto M. Chemometrics in kinetic analysis. *Analyst* 1990; **115**: 685–688.
3. Crouch SR. Trends in kinetic methods of analysis. *Anal. Chim. Acta* 1993; **283**: 453–470.
4. Kaufman D, Sterner C, Masek B, Svenningsen R, Samuelson G. An NMR kinetics experiment. *J. Chem. Educ.* 1982; **59**: 885–886.
5. Chrastil J. Determination of the first order consecutive reaction rate constants from final product. *Comput. Chem.* 1988; **12**: 289–292.
6. Chrastil J. Determination of the first-order consecutive reversible reaction kinetics. *Comput. Chem.* 1993; **17**: 103–106.
7. Sylvestre EA, Lawton WH, Maggio MS. Curve resolution using a postulated chemical reaction. *Technometrics* 1974; **16**: 353–368.
8. Mayes DM, Kelly JJ, Callis JB. Non-invasive monitoring of a two-step sequential chemical reaction with shortwave near-infrared spectroscopy, In *Near Infra-red Spectroscopy: Bridging the Gap between Data Analysis and NIR Applications*, Hildrum KI, Isaksson T, Naes T, Tandberg A (eds). Ellis Horwood: Chichester, 1992; 377–387.
9. Bugnon P, Chottard J-C, Jestin J-L, Jung B, Laurency G, Maeder M, Merbach AE, Zuberbühler AD.

- Second-order globalisation for the determination of activation parameters in kinetics. *Anal. Chim. Acta* 1994; **298**: 193–201.
10. Tam KY, Chau FT. Multivariate study of kinetic data for a two-step consecutive reaction using target factor analysis. *Chemometrics Intell. Lab. Syst.* 1994; **25**: 25–42.
  11. Bijlsma S, Louwerse DJ, Smilde AK. Rapid estimation of rate constants of batch processes using on-line SW-NIR. *AIChE J.* 1998; **44**: 2713–2723.
  12. Chau F-T, Mok K-W. Multiwavelength analysis for a first-order consecutive reaction. *Comput. Chem.* 1992; **16**: 239–242.
  13. Bijlsma S, Smilde AK. Application of curve resolution based methods to kinetic data. *Anal. Chim. Acta* 1999; **396**: 231–240.
  14. Mok K-W, Chau F-T. Application of the extended Kalman filter for analysis of a consecutive first-order reaction. *Trends Anal. Chem.* 1996; **15**: 170–174.
  15. Bijlsma S, Louwerse DJ, Windig W, Smilde AK. Rapid estimation of rate constants using on-line SW-NIR and trilinear models. *Anal. Chim. Acta* 1998; **376**: 339–355.
  16. Bijlsma S, Louwerse DJ, Smilde AK. Estimating rate constants and pure UV-VIS spectra of a two-step reaction using trilinear models. *J. Chemometrics* 1999; **13**: 1–19.
  17. Windig W, Antalek B, Sorriero LJ, Bijlsma S, Louwerse DJ, Smilde AK. Applications and new developments of the direct exponential curve resolution algorithm (DECRA). Examples of spectra and magnetic resonance images. *J. Chemometrics* 1999; **13**: 95–110.
  18. Maeder M, Molloy KJ, Schumacher MM. Analysis of non-isothermal kinetic measurements. *Anal. Chim. Acta* 1997; **337**: 73–81.
  19. Furusjö E, Danielsson L-G. A method for the determination of reaction mechanisms and rate constants from two-way spectroscopic data. *Anal. Chim. Acta* 1998; **373**: 83–94.
  20. Molloy KJ, Maeder M, Schumacher MM. Hard modelling of spectroscopic measurements. Applications in non-ideal industrial reaction systems. *Chemometrics Intell. Lab. Syst.* 1999; **46**: 221–230.
  21. Lawton WH, Sylvestre EA. Self modeling curve resolution. *Technometrics* 1971; **13**: 617–633.
  22. Tauler R, Fleming S, Kowalski BR. Multivariate curve resolution applied to spectral data from multiple runs of an industrial process. *Anal. Chem.* 1993; **65**: 2040–2047.
  23. Lacorte S, Barceló D, Tauler R. Determination of traces of herbicide mixtures in water by on-line solid-phase extraction followed by liquid chromatography with diode-array detection and multivariate self-modelling curve resolution. *J. Chromatogr. A* 1995; **697**: 345–355.
  24. Tauler R, Izquierdo-Ridora A, Gargallo R, Casassas E. Application of a new multivariate curve resolution procedure to the simultaneous analysis of several spectroscopic titrations of the copper (II)-polyinosinic acid system. *Chemometrics Intell. Lab. Syst.* 1995; **27**: 163–174.
  25. Shrager RI, Hendler RW. Titration of individual components in a mixture with resolution of difference spectra, pKs, and redox transitions. *Anal. Chem.* 1982; **54**: 1147–1152.
  26. Frans SD, Harris JM. Reiterative least-squares spectral resolution of organic acid/base mixtures. *Anal. Chem.* 1984; **56**: 466–470.
  27. Frans SD, Harris JM. Least squares singular value decomposition for the resolution of pKs and spectra from organic acid/base mixtures. *Anal. Chem.* 1985; **57**: 1718–1721.
  28. Shrager RI. Chemical transitions measured by spectra and resolved using singular value decomposition. *Chemometrics Intell. Lab. Syst.* 1986; **1**: 59–70.
  29. Antalek B, Windig W. Generalized rank annihilation method applied to a single multicomponent pulsed gradient spin echo NMR data set. *J. Am. Chem. Soc.* 1996; **118**: 10331–10332.
  30. Windig W, Antalek B. Direct exponential curve resolution algorithm (DECRA): a novel application of the generalized rank annihilation method for a single spectral mixture data set with exponentially decaying contribution profiles. *Chemometrics Intell. Lab. Syst.* 1997; **37**: 241–254.
  31. Windig W, Hornak JP, Antalek B. Multivariate image analysis of magnetic resonance images with the direct exponential curve resolution algorithm (DECRA). Part 1: Algorithm and model study. *J. Magn. Reson.* 1998; **132**: 298–306.
  32. Antalek B, Hornak JP, Windig W. Multivariate image analysis of magnetic resonance images with the direct exponential curve resolution algorithm (DECRA). Part 2: Application to human brain images. *J. Magn. Reson.* 1998; **132**: 307–315.
  33. Wilson B, Sanchez E, Kowalski BR. An improved algorithm for the generalized rank annihilation method. *J. Chemometrics* 1989; **3**: 493–498.
  34. Seber GAF, Wild CJ. *Nonlinear Regression*. Wiley: New York, 1989.
  35. Harshman RA, Lundy ME. PARAFAC: parallel factor analysis. *Comput. Statist. Data Anal.* 1994; **18**: 39–72.

36. Smilde AK. Three-way analyses problems and prospects. *Chemometrics Intell. Lab. Syst.* 1992; **15**: 143–157.
37. Shao J, Tu D. *The Jackknife and Bootstrap*. Springer: New York, 1995.
38. Booksh KS, Lin Z, Wang Z, Kowalski BR. Extension of trilinear decomposition method with an application to the flow probe sensor. *Anal. Chem.* 1994, **66**: 2561–2569.
39. Hairfield EM, Moomaw EW, Tamburri RA, Vigil RA. The epoxidation of 2,5-di-tert-butyl-1,4-benzoquinone. *J. Chem. Educ.* 1985, **62**: 175–177.
40. Bisby RH, Thomas EW. Kinetic analysis by the method of nonlinear least squares. *J. Chem. Educ.* 1986, **63**: 990–992.
41. Burns DA, Ciurczak EW. *Handbook of Near-infrared Analysis*. Dekker: New York, 1992.
42. Bro R, de Jong S. A fast non-negativity-constrained least squares algorithm. *J. Chemometrics* 1997, **11**: 393–402.
43. Kiers HAL. Weighted least squares fitting using ordinary least squares algorithms. *Psychometrika* 1997, **62**: 251–266.
44. Kruskal JB. Rank, decomposition, and uniqueness for 3-way and n-way arrays. In *Multiway Data Analysis*, Coppi R, Bolasco S (eds). Elsevier: Amsterdam, 1989; 7–18.
45. Rao CR, Mitra S. *Generalized Inverse of Matrices and Its Applications*. Wiley: New York, 1971.
46. Savitzky A, Golay MJE. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* 1964, **36**: 1627–1639.
47. Faber NM, Buydens LMC, Kateman G. Generalized rank annihilation method. II: Bias and variance in the estimated eigenvalues. *J. Chemometrics* 1994, **8**: 181–203.