

# Characterization of Unimodality Constraint as an Effective Chemistry-based Condition in Resolving of Chemical Processes Data

Somaye Vali Zade<sup>a,b</sup>, Klaus Neymeyr<sup>b,c</sup>, Hamid Abdollahi<sup>a</sup>, Mathias Sawall<sup>b</sup>

<sup>a</sup>Faculty of Chemistry, Institute for Advanced Studies in Basic Sciences, 45195-1159 Zanjan, Iran

<sup>b</sup>Universität Rostock, Institut für Mathematik, Ulmenstrasse 69, 18057 Rostock, Germany

<sup>c</sup>Leibniz-Institut für Katalyse, Albert-Einstein-Strasse 29a, 18059 Rostock, Germany

---

## Abstract

Multivariate curve resolution (MCR) techniques are powerful methods for qualitative and quantitative analyses of spectral data. However, MCR methods suffer from non-unique solutions - instead of a unique solution usually a continuum of possible pure component decomposition exists. This ambiguity is covered in a low-dimensional way by the so-called area feasible solutions (AFS). In the quantitative purpose, the rotational ambiguity has a negative effect on the accuracy of the results. By applying additional constraints on the profiles the range of feasible solutions can considerably improve the reliability of quantitative results. Another approach to analyze the size of the continuum of possible pure component profiles is to compute the extrema of the signal contribution function (SCF). The SCF can also be applied under additional constraints.

In this study we analyze the restriction of the AFS and the change of the SCF-extrema under additional unimodality constraints. An interesting effect is stated for two examples namely that the unimodality constraint can divide a single subset of the AFS into two parts. In such cases the SCF cannot resolve the ambiguity completely. The effect of the unimodality constraints has been systematically investigated on two- and three-component simulation data sets. This phenomenon are also studied for an experimental data set related to the spectrophotometric titration of a mixture of ligand (3, 7-diazanonanedioic acid diamide) and  $\text{Cu}^{2+}$  with varying degrees of acidity. Imposing of unimodality constraints as an effective chemically meaningful limitation on the feasible solutions of MCR methods is strongly recommended.

**Keywords:** multivariate curve resolution, self-modeling curve resolution, area of feasible solutions, FACPAC, MCR-BANDS, unimodality constraint.

---

## 1. Introduction

Often in analytical chemistry, the analysis of complex samples in terms of identifying the underlying pure components and their concentrations is by no means trivial. Chromatographic systems are focussed on the analysis of trace amounts of components in complex samples. The evolution of new multivariate curve resolution (MCR) techniques for separation, identification, and quantification of pure component profiles in data matrices from chromatographic samples has become one of the most important developments in analytical chemistry in recent years. However, one of the main challenges in chromatography is to handle and analyze the enormous amounts of data obtained in these cases in order to interpret the results [20].

Chemometric methods are powerful tools to analyze multivariate data from modern analytical instruments. Some of these methods, in particular multivariate curve resolution methods (MCR) [8], have proved their potential for analyzing and interpreting multivariate chromatographic data in terms of obtaining pure component profiles of interest. Also, a combination of MCR methods and chromatographic systems offer the possibility of resolution, analysis and quantitation of multiple components in unknown mixtures without their previous chemical or physical separations.

The applications of MCR methods in second-order calibration have been illustrated in different publications [5, 37] and it is proved to be a powerful technique for quantification of complex mixtures. However, the main limitation of all MCR methods is the rotational ambiguity which is known and documented since the very beginning of chemometrics [16]. The reliability and accuracy of quantitative MCR results strongly depend on the rotational ambiguity of factorizations of the system under study.

The underlying bilinear model for these methods is the Lambert-Beer law [7, 16–18, 26]. Let a series of spectra be given with a number of  $k$  spectra taken at  $n$  frequencies. Thus the data is considered in a time  $\times$  frequencies-grid. The mixed spectra are stored in a spectral data matrix  $D \in \mathbb{R}^{k \times n}$ . According to the Lambert-Beer law in matrix notation  $D$

can (approximately) be factorized into a product of a concentration matrix  $C \in \mathbb{R}^{k \times s}$  and a matrix  $S^T \in \mathbb{R}^{s \times n}$  of pure component spectra

$$D = CS^T. \quad (1)$$

Therein  $s$  equals the number of chemical components. The profiles are stored column-wise in the matrices  $C$  and  $S$ . The goal of MCR techniques is to extract the unknown pure component information of the underlying chemical system in form of  $C$  and  $S$  from  $D$ . This means that a nonnegative matrix factorization of  $D$  is to be determined. All approaches to compute decompositions as in (1) have the intrinsic drawback that not only one decomposition exists but continua of possible factorizations. This fact is known under the keyword of rotational ambiguity of the solution/factorization [3, 41]. The presence of rotational ambiguities and non-unique solutions decreases the reliability of a possible factorization and makes their assessment difficult.

The Area of Feasible Solutions (AFS, [7, 11, 12, 25, 27, 29]) and the signal contribution function (SCF, [10, 19, 38]) are two approaches to determine and analyze the range of feasible solutions. The AFS is a low-dimensional representation of all profiles that can be part of all possible factorizations according to (1) with  $C, S \geq 0$ . Since the model is bilinear two AFS-sets exist, namely the concentration AFS and the spectral AFS. The polygon inflation algorithm [29, 31] as implemented in FACPAC is an efficient and robust method for the computation of the AFS-sets. A further important approach to solve the MCR problem is to compute local extrema of the SCF. For each chemical component band boundaries can be computed that limit the bands of possible concentration profiles and also the band of spectral profiles. This enables an estimation of the degree of freedom for the profiles of each pure component. The method is implemented in the MCR-BANDS software.

In order to reduce the impact of the rotational ambiguity and hopefully to obtain a unique (true) solution, additional information about the chemical system can be applied in the form of proper constraints. This results in a reduction of the set of all nonnegative factorizations to a set of all nonnegative and meaningful solutions in terms of the certain constraint. Constraints can be applied in the form of chemical conditions or in the form of restrictions on the profile shape (we call this a mathematical constraint). The resolved profiles should respect these constraints. The rigorous application of mathematical constraints may lead to incorrect solutions; Akbari et al. [4] have shown that some mathematical constraints may not be in agreement with chemical information.

Physico-chemical meaningful constraints are, e. g., nonnegativity of concentrations and absorption values and unimodality of elution profiles. The unimodality constraint allows the presence of only one maximum per profile [38]. This condition is fulfilled by many peak-shaped concentration profiles, like chromatograms or some types of reaction profiles. It is important to note that this constraint cannot only be applied to peaks, but also to profiles that have a constant maximum (plateau) or a decreasing tendency such as the most protonated and deprotonated species in an acid-base titration. In these cases the profiles behave monotonously and monotonicity constraints can be applied. Monotonicity is a restricted mode of unimodality and the difference between unimodality and monotonicity constraints is well illustrated in [40].

Unimodality has been used as a physico-chemical based information on concentration profiles in different kinetics and equilibria chemical processes as well as for response profiles in spectrofluorimetry emission and electrochemical responses [2]. Tauler et al. used MCR-ALS incorporating the unimodality constraints for the resolution and quantitation of mixture of two co-eluted compounds in HPLC-DAD [39]. In 2009, Beyramisoltan et al. suggested a simple procedure for implementing the unimodality constraint in the Lawton-Sylvestre method [6]. Sawall et al. investigated the impact of soft unimodality constraints to reduce the AFS-sets [35]. Vali Zade et al. investigated the difference between monotonicity and unimodality constraints [40].

Despite the widespread application of the unimodality constraints, their impact on the reduction of the rotational ambiguity and their consequent increasing of the reliability of MCR results, a systematic investigation of the impact of the unimodality constraint has not been studied so far. In this contribution, we investigate the effect of the unimodality constraint on reducing the AFS and also the high impact of unimodality on increasing the accuracy of the results. As an interesting point we analyze and discuss the effect that AFS-subset can split up into two subsets if additional unimodality constraints are applied in Sec. 4. We also compare the unimodality-restricted MCR-BANDS solutions with the unimodality restricted AFS-sets. The aim is to get a deeper understanding of MCR methods under unimodality constraints. The information from this study can be very important in the application of MCR methods in analyzing different chemical data sets.

## 2. Theory

### 2.1. The area feasible solutions

AFS is a very powerful tool for investigating the effect of imposing information in MCR methods on the accuracy of the results. Together with many MCR-approaches the AFS representation of the rotational ambiguity is based

on forming the pure component factors  $C$  and  $S$  by using a singular value decomposition (SVD, [13]) of  $D$ . An SVD  $D = U\Sigma V^T$  decomposes  $D$  into a matrix of left singular vectors  $U$ , a matrix of right singular vectors  $V$  and a (rectangular) diagonal matrix  $\Sigma$  with the singular values on the diagonal. If  $D$  contains absorptions of  $s$  components, then a truncated SVD ( $U$  and  $V$  are reduced to their first  $s$  columns and  $\Sigma$  is reduced to its leading  $s \times s$  submatrix) can be used in order to reconstruct  $C$  and  $S$  as

$$C = U\Sigma T^{-1} \quad S^T = TV^T$$

with a proper regular matrix  $T \in \mathbb{R}^{s \times s}$ , see e.g. [16, 17].

In order to analyze the ambiguity of the factorization one is interested in the sets of all possible columns for  $C$  and  $S$  that can be part of a nonnegative matrix factorization of  $D$ . The AFS-sets are low-dimensional representations of the possible columns of  $C$  and  $S$ . Since the columns of  $C$  and  $S$  can simultaneously be permuted without changing the chemical information content, one can justify to consider only the set of all possible first rows of  $T$ . Furthermore, the scaling of the columns of these matrices cannot be determined without additional information (as calibration). Therefore and without loss of generality the pure component spectra are calibrated in a way that all matrix elements in the first column of  $T$  are equal to 1. Hence we get

$$T = \begin{pmatrix} 1 & x_1 & \cdots & x_{s-1} \\ 1 & & & \\ \vdots & & \mathbf{W} & \\ 1 & & & \end{pmatrix}. \quad (2)$$

With  $T$  being of the form of (2) the spectral factor AFS reads

$$\mathcal{M}_S = \left\{ x \in \mathbb{R}^{s-1} : \text{exists } W \in \mathbb{R}^{(s-1) \times (s-1)} \text{ with } \text{rank}(T) = s \text{ and } C, S \geq 0 \right\}.$$

The concentration factor AFS  $\mathcal{M}_C$  can be defined in an analogous way. Here we use the definition from (5) of [28]. For the computation of the AFS-sets several methods are known: For chemical systems with two components an analytical solution for determining feasible regions for the MCR solutions was introduced in 1971 [16]. For these two-component systems the boundaries of the feasible intervals can be computed directly [16, 27] or by using a numerical grid-search optimization [41]. For three-component systems a geometric construction [7, 15, 25, 28] as well as several numerical approximation methods [12, 29, 32, 33] are available. See also [30] for a hybrid method.

The geometric constructions of the AFS-sets (Borgen plots) are based on properly constructed triangles that are located between the inner polygon and the outer polygon [7, 15, 25, 28]. The outer polygon contains all low-dimensional representations that lead to nonnegative profiles  $S(:, 1)$ . The faces (sides) of the outer polygon are the boundaries in the abstract space between sections representing exclusively nonnegative profiles and sections representing partly-negative profiles. The transformed profiles from the faces of the outer polygon have at least one zero value. The inner polygon is the convex hull of the low-dimensional representations of the columns (respectively rows for the dual AFS) of  $D$  [24]. The AFS-sets contain the low-dimensional representations of all possible profiles for the curve resolution problem and are located between the inner and the outer polygons. Typically the AFS-sets consist of several separated subsets (sometimes called segments). Sometimes the AFS is a connected set with a hole around the origin. See [11, 12, 27–29] for more details on the AFS-approach as well as the outer polygons  $\mathcal{F}_S, \mathcal{F}_C$  and the inner polygons  $\mathcal{I}_S, \mathcal{I}_C$ .

## 2.2. AFS computation

In this paper the AFS-sets for two-component resp. three-component systems are computed by grid search [41] respectively the polygon inflation algorithm [29, 31]. The idea of grid search is to discretize the outer polygon by a fine grid and to check whether the nodes of the grid belong to  $\mathcal{M}_S$  respectively  $\mathcal{M}_C$ . The idea of the polygon inflation methods is to approximate the boundaries of (all subsets of the) AFS by sequences of adaptively refined polygons. In the case of an AFS set containing a hole, this hole boundary is approximated similarly. The inflation algorithm subdivides the edges of the polygons recursively and adds new vertices that are located on the boundary of the AFS-sets. A local error estimation is used to control the polygon refinement process. Other AFS computation routines are the geometric construction [7, 25, 28], triangle enclosure [12] and particle swarm approximation [36].

The grid search approach as well as the polygon inflation algorithm are based on a cost function that check whether points  $x \in \mathbb{R}^{s-1}$  belong to the AFS or not. The numerical evaluation is based on an optimization process that can include additional constraints as parts of the cost function. See e.g. [12, 29] for the possible cost functions without additional constraints and [21, 34, 40] how additional constraints as monotonicity or unimodality can be used in a cost function.

### 2.3. The signal contribution function, its extrema and the software MCR-BANDS

Based on a certain factorization  $D = CS^T$  the SCF is defined for the  $i$ th chemical component as

$$\text{SCF}_i = \frac{\left\| \left( CR^{-1}(:, i) \right) \left( R(i, :) S^T \right) \right\|_F}{\|D\|_F} \quad (3)$$

with a regular matrix  $R \in \mathbb{R}^{s \times s}$ . Thus the SCF returns the ratio of the individual signal contribution of component  $i$  to the full signal of all components. The elements of the matrix  $R$  are subject of optimizations with nonlinear constraints in order to maximize respectively to minimize the SCF for each component.

The idea to minimize or to maximize the SCF has been proposed in order to estimate the degree of the rotational ambiguity linked to a particular MCR solution. A local extremum of the SCF defines a pair of a spectrum and a concentration profile fulfilling proper constraints and giving a relative maximal or minimal signal contribution over the whole signal from all components. The difference between the extremal profiles indicates the rotational ambiguity for the associated profile. In case of the presence of a large rotational ambiguities, the differences will also be large. If no rotational ambiguity exists for one profile, then the difference will be zero. See [1, 9, 10, 14, 23, 38] for more details about this procedure and its implementation. Recently, a complete mathematical proof has been given in [19] that for two-component systems the representations of the extrema of the SCF are always on the boundary of the AFS. The MCR-BANDS software with a graphical user interface for MATLAB is freely available under the MCR-ALS-web-page <http://www.mcrals.info>.

### 2.4. Unimodality constraints

A profile is called unimodal if it has only one maximum and is monotone increasing before and monotone decreasing after this maximum. In terms of the reconstruction problem according (1) unimodality constraints reduce the set of all pure component decompositions to a set of factorizations with only unimodal profiles. Unimodality is very important constraint for improving the accuracy of data from second-order chromatography and chemical process monitoring - two common issues in analytical chemistry. So, our results about explanation on the nature of this property will be very useful for analytical chemists. There are some primary studies on the restrictive effect of unimodality constraints to the AFS has already been shown in [6, 21, 34].

In terms of the AFS unimodality constraints reduce the outer polygon  $\mathcal{F}_C$  respectively  $\mathcal{F}_S$ , depending for which factor unimodality is applied, to a subset of elements that represents only unimodal profiles. Since the AFS is a subset of the outer polygon, areas that do not represent unimodal profiles are cut away. For details on further reduction steps for the two AFS-sets see Sec. 3.2.

### 2.5. Splitting effect under unimodality constraints

The additional application of unimodality constraints can result in a splitting of the original AFS-subset. As far as the authors know such an effect has not been analyzed yet. The splitting phenomenon is not restricted to artificially constructed model problems but can be observed even for experimental spectral data. In Sec. 4 two data sets for  $s = 2$  and  $s = 3$  are discussed for which certain AFS-subsets split up under unimodality constraints. The effect is independent from the computational method for calculating the feasible solutions and is also not related to the underlying optimization processes. We think that this observation might be interesting for analytical chemists. For split AFS-sets the SCF-approach suffers from the representation by only two extreme points that cannot resolve a possible gap. In the following two sections we discuss the two cases of non-splitting and those of splitting AFS-subsets.

## 3. Results for non-split subsets

In this section, we compare SCF-extrema computed by MCR-BANDS with the AFS-sets. A two- and a three-component model problem are analyzed. Nonnegativity constraints are applied to each factor. Unimodality constraints are applied only to factor  $C$ . For the model problems in this section the AFS-subsets do not split up if unimodality constraints are applied.

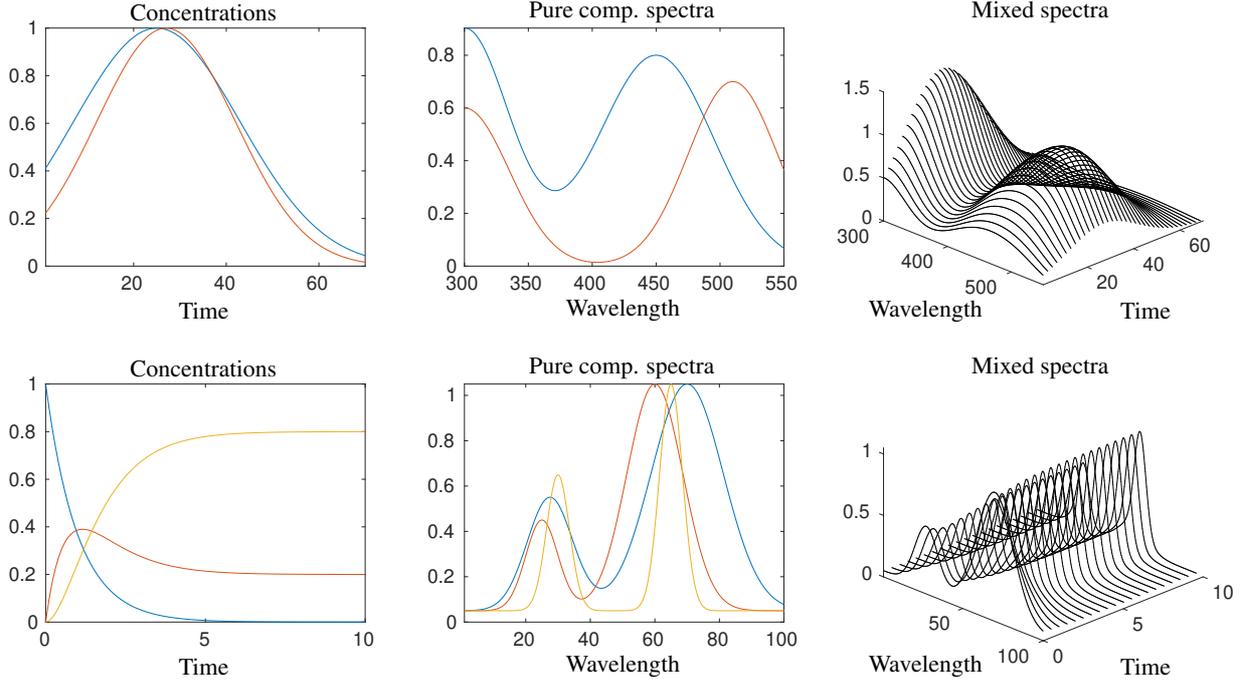


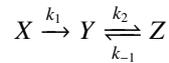
Figure 1: A two- and a three-component model problem as used in Sec. 3. The left panel shows the simulated concentration profiles, the central panel shows the simulated spectral profiles and the right panel shows the mixed data (only every 2nd respectively every 5th profile is plotted).

### 3.1. Two-component system

The two-component data set is based on a chromatography model problem. The data contains  $k = 70$  spectra for  $n = 251$  channels. The concentration profiles, the pure component spectra as well as the mixed spectra are shown in the upper row of plots in Fig. 1. The AFS-sets and the SCF-extrema are computed under nonnegativity constraints and also under nonnegativity plus unimodality constraints for factor  $C$ . The AFS-sets are computed by grid search and the SCF-extrema are computed by MCR-BANDS. The results are presented in Fig. 2. The feasible intervals (the pendant of the AFS for two-component systems) are plotted as lower lines. These intervals for nonnegativity and unimodality constraints are plotted as upper lines. The SCF-extrema for using nonnegativity constraints are attained in endpoints of these intervals. This is also the case if additionally unimodality constraints are used. The impact of the unimodality constraint is obvious and the indirect restriction for the factor  $S$  is caused by duality. The feasible profiles under nonnegativity and unimodality constraints are presented in Fig. 2.

### 3.2. Three-component model problem

Next we study the model problem for the reaction equations



with  $k_1 = 1$ ,  $k_2 = 1$  and  $k_{-1} = 0.25$ . The factor  $C$  is computed for the initial concentrations  $c_x(0) = 1$ ,  $c_y(0) = c_z(0) = 0$  and for equidistant points in time with  $k = 100$  nodes and  $t \in [0, 10]$ . The pure component spectra are evaluated on an equidistant frequency grid with  $n = 250$  nodes for  $\lambda \in [0, 100]$  and

$$\begin{aligned} s_x(\lambda) &= 0.5 \exp\left(-(\lambda - 27.5)^2/100\right) + \exp\left(-(\lambda - 70)^2/250\right), \\ s_y(\lambda) &= 0.4 \exp\left(-(\lambda - 25)^2/50\right) + \exp\left(-(\lambda - 60)^2/150\right), \\ s_z(\lambda) &= 0.6 \exp\left(-(\lambda - 30)^2/25\right) + \exp\left(-(\lambda - 65)^2/25\right). \end{aligned}$$

The concentration profiles and the pure component spectra as well as the mixed profiles are shown in Fig. 1 (2nd row).

First we analyze the impact of unimodality constraints to the outer polygon in the  $U$ -space, namely  $\mathcal{F}_C$ , see (9) in [27] for the formal definition of  $\mathcal{F}_C$ . Therefore we compute the areas of  $\mathcal{F}_C$  for which the transformed concentration profiles are nonnegative and unimodal respectively nonnegative and not unimodal [40]. Fig. 3 (left) shows the unimodality/non-unimodality areas inside the outer polygon  $\mathcal{F}_C$  for the model data. Points in the red colored area

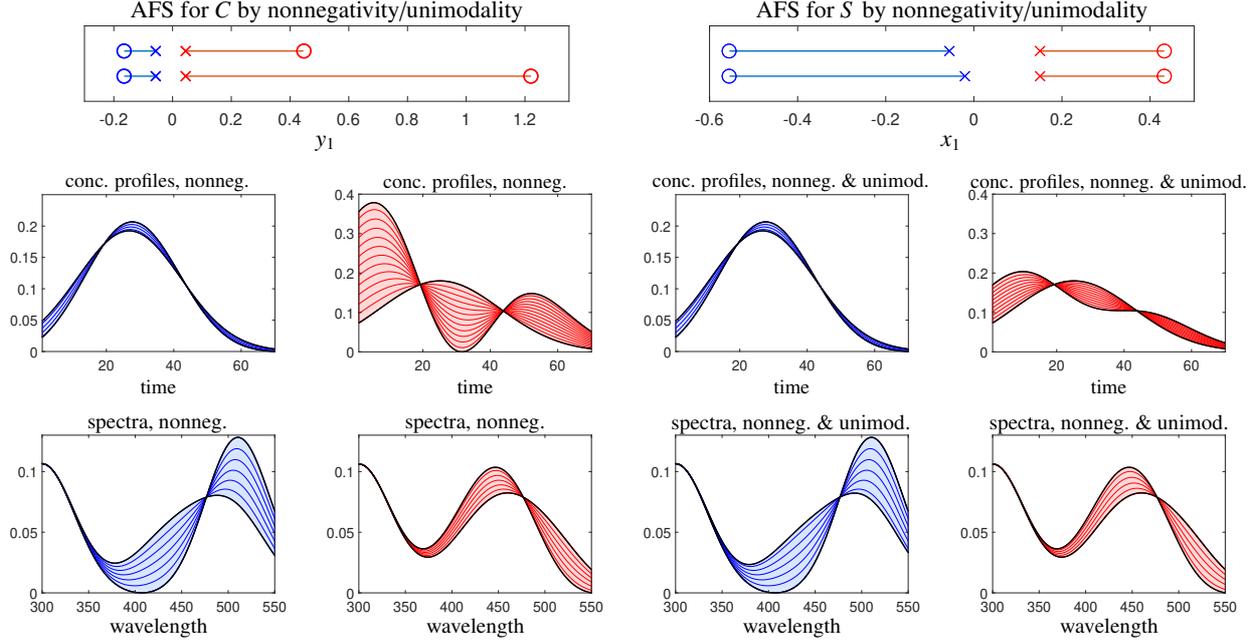


Figure 2: The AFS-sets as well as the SCF-extrema with and without additional unimodality constraints (top) as well as the associated feasible and extremal profiles (2nd and 3rd row) for the data set from Sec. 3.1. In the AFS-plots (feasible intervals) the lower lines are the results which are gained only under nonnegativity constraints. The upper lines are attained if additionally unimodality constraints for factor  $C$  are applied. The SCF-minima respectively -maxima are marked by  $\circ$  resp.  $\times$ . All extrema are attained on the boundary (endpoints) of the feasible intervals. The feasible profiles in which the SCF takes its extrema are plotted as black lines. These extremal profiles enclose the bands of all feasible profiles.

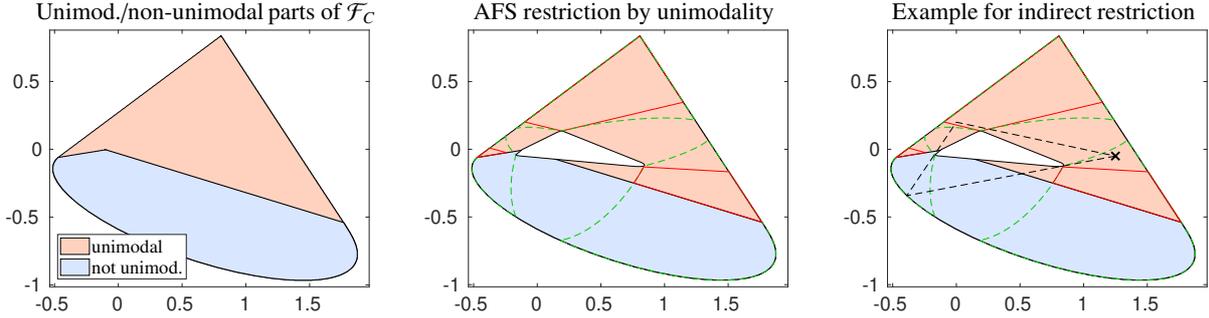


Figure 3: Unimodal and non-unimodal areas of the outer polygon in concentration space  $\mathcal{F}_C$  (left), their direct and indirect restrictions (center) and an example for the indirect restriction (right) for the model problem from Sec. 3.2. Left: Points in the red colored area represent unimodal profiles and points in the blue colored area represent non-unimodal profiles. The red lines are the boundaries of the AFS-sets under nonnegativity and unimodality constraints. The white area represents the inner polygon  $\mathcal{I}_C$ , see [27]. Right: An example for the indirect restriction by unimodality constraints. A point (black cross) that represents a nonnegative and unimodal profile is selected. However, it does not belong to the unimodality-AFS, since no triangle (black dashed lines) with vertices in the red area exists that encloses the inner polygon.

result in unimodal concentration profiles and points in the blue colored area are related to unimodal concentration profiles.

The reduction of  $\mathcal{F}_C$  has a direct impact in terms of a reduction of the feasible regions in the  $U$ -space as well as an indirect impact, namely that certain simplex constructions (Borgen triangle) are not possible any more. We point out the important fact: Although a point belongs to the nonnegativity AFS and represents a unimodal profiles, the point does not necessarily belongs to the unimodality restricted AFS. In the right plot of Fig. 3 we present this observation for a certain solution. A certain element (black  $\times$ ) of the nonnegativity AFS (green dashed lines) is selected that represents a unimodal profile. But no other two vertices that represent unimodal profiles exist (see the black dashed lines) such that the triangle encloses the inner polygon (feasible triangle). So this point is excluded from the unimodality restricted AFS.

The results for the model problem are presented in Fig. 4. The AFS-sets with and without additional unimodality constraints are computed by polygon inflation and the SCF-extrema are computed by MCR-BANDS. For the SCF-extrema we get six extremal concentration profiles and six extremal pure component spectra because two solutions

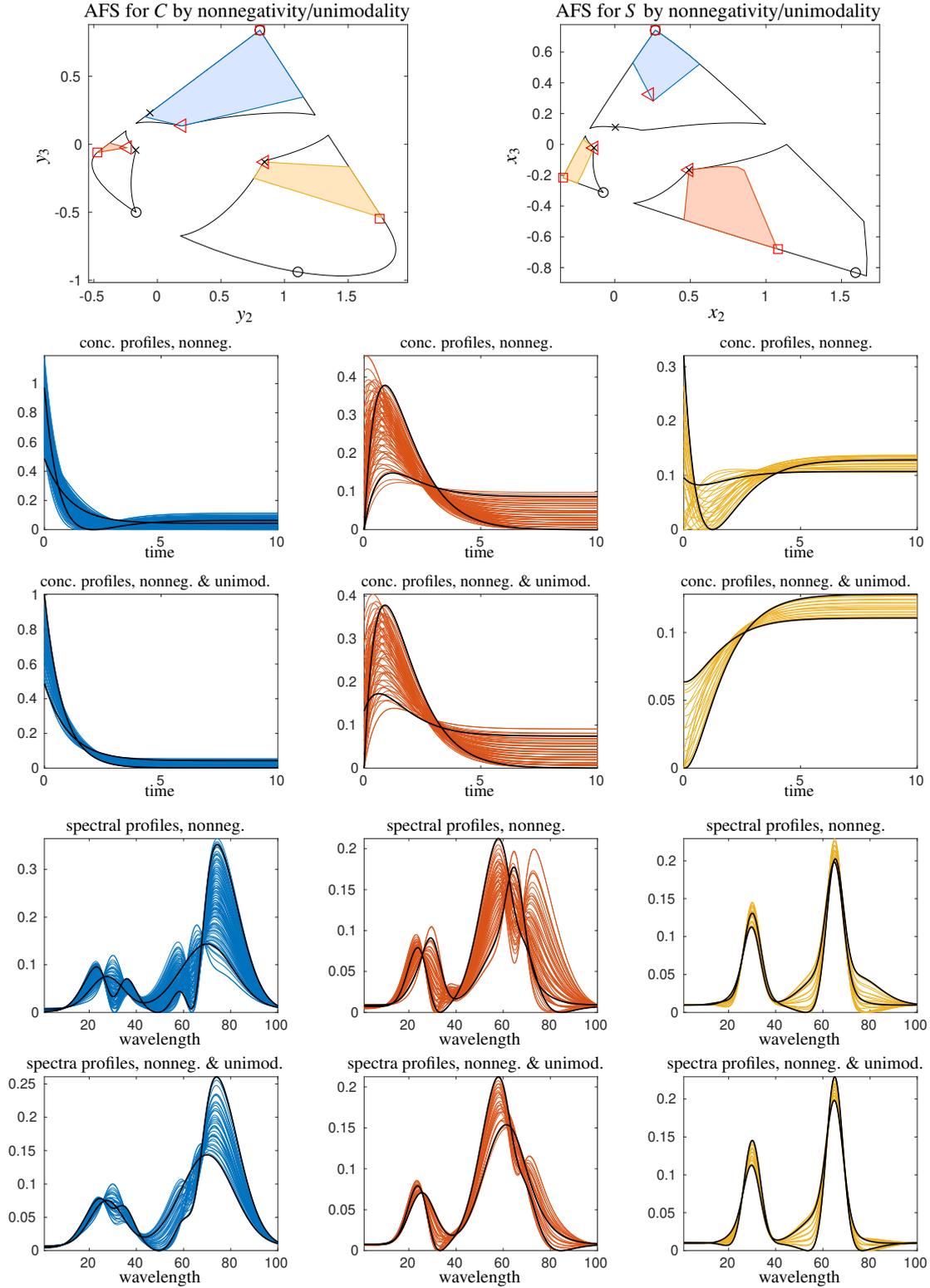


Figure 4: Results for the analysis of the rotational ambiguity for the data set of Sec. 3.2 by the AFS and the SCF-extrema. Top: The AFS-sets and the SCF-extrema with and without additional unimodality constraints. The boundaries of the nonnegativity AFS are plotted in black, the unimodality AFS is colored. The SCF-extrema using only nonnegativity constraints are plotted in black, the unimodality restricted extrema are plotted in red. For both approaches the restrictive effect of the unimodality constraints is obvious. For each subset the SCF-extrema are on the AFS-boundary but the SCF-solutions cannot represent the whole ambiguity as the AFS. Rows 2-5: The feasible profiles obtained by the AFS (colored profiles) and the SCF-extrema (black lines) corresponding to the single components ( $X$ ,  $Y$  and  $Z$ ). Applied constraints: nonnegative for  $C$  and  $S$  (2nd and 4th row), nonnegativity for  $C$  and  $S$  and unimodality for  $C$  (3rd and 5th row).

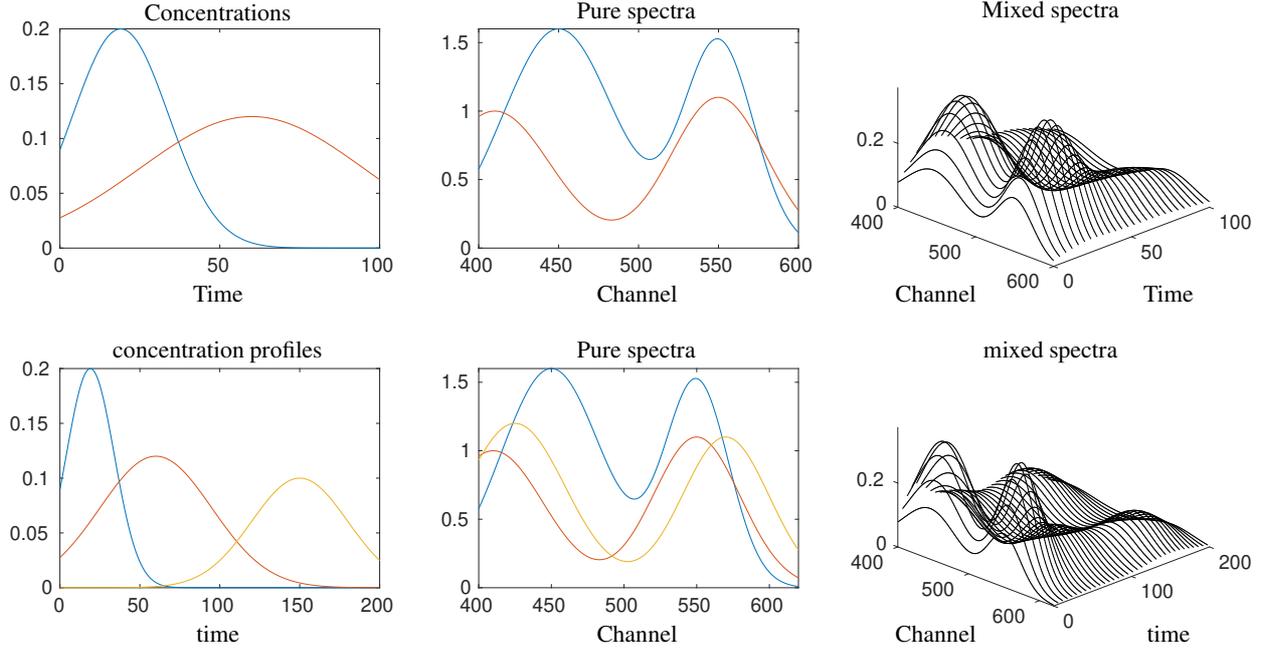


Figure 5: A two- and a three-component model problem as used in Sec. 4. The left panel shows the simulated concentration profiles, the central panel shows the simulated spectral profiles and the right panel shows the mixed data (only every 4th resp. 6th profile is plotted).

are computed for every profile. Three of the profiles correspond to local maxima of the SCF and also three profiles correspond to local minima of the SCF. The impact of the unimodality constraints is obvious for both approaches. The SCF-extrema that are outside of the unimodality-AFS shift to the boundary of the unimodality-AFS if unimodality constraints are applied. The two approaches are not in mutual conflict. The feasible concentration profiles and pure component spectra are also presented in Fig. 4. The profiles by the SCF-approach are limits in terms of the values of the SCF but not in term of their graphical representation embedding all feasible profiles.

#### 4. Results for split subsets

In this section we present results for two data sets for which certain AFS-subsets split up under unimodality constraints. A two- and a three-component system are considered. If only nonnegativity constraints are applied to these data sets, then the AFS consists of three clearly separated subsets (and each of these three subsets is a connected set). If additionally unimodality constraints are applied, then for each of the two data sets one AFS subset in the concentration factor AFS and also one AFS subset in the spectral AFS splits up into two subsets. It is important to remark that such a separation is not uncommon and is observed not only for model problems. The two data sets as well as the pure factors are shown in Fig. 5.

In order to be able to compute split AFS-sets for divided subsets AFS-computation methods as polygon inflation [29, 31] or triangle enclosure [12] need additional algorithmic elements. However, the limitations of polygon inflation and triangle enclosure are not of conceptual nature but are a consequence of the assumptions on the number of isolated subsets of an AFS. As the classical grid search algorithm makes no assumptions on this number of subsets of an AFS, it works without modifications but it needs a relatively high resolution. Since the SCF-extrema are by their construction only two points per AFS-subset they cannot represent a possible gap.

##### 4.1. A two-component model problem with a split subset

The concentration profiles, the pure component spectra as well as the mixed spectra of the two-component model problem are shown in Fig. 5 (top). The data matrix consists of  $k = 70$  spectra for  $n = 251$  wavelengths.

The AFS-sets as well as the SCF-extrema with and without additional unimodality constraints are presented in Fig. 6. The AFS by using nonnegativity constraints (only for factor  $C$ ) is plotted by the lower lines. The AFS by using nonnegativity and unimodality constraints is plotted by the upper lines. The SCF-extrema under nonnegativity constraints for two-component systems are attained in the endpoints of the AFS-intervals as expected, see [19]. Furthermore, if additionally unimodality constraints are applied, then one subset of the concentration factor AFS splits up into two parts. As an indirect effect, also one subset of the spectral AFS splits up. The subdivisions of the AFS cannot



Figure 6: The AFS-sets and the extrema of the SCF with (top) and without (bottom) additional unimodality constraints for the concentration profiles for the data set as introduced in Sec. 4.1. If unimodality constraints are applied, then one subset of the concentration factor AFS splits up into two parts. The same does one subset of the spectral AFS. These separations cannot be represented completely by the extrema of the SCF. Furthermore as an interesting indirect effect one extremum of the SCF is not located in an endpoint of the interval since the gap in the dual AFS is not bridged over in the optimization process.

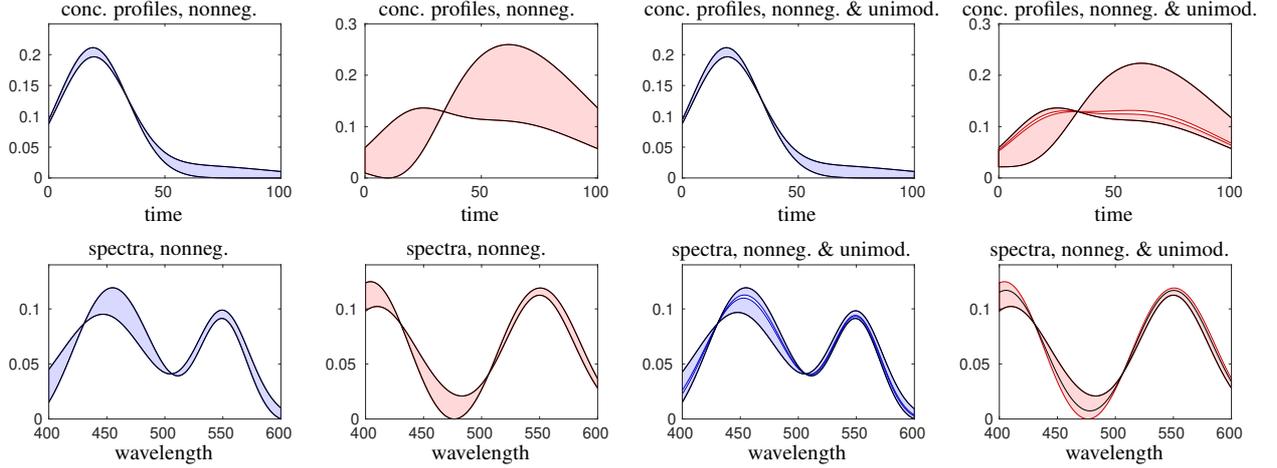


Figure 7: Feasible profiles for the factors  $C$  and  $S^T$  obtained by grid search (colored profiles) and MCR-BANDS (black profiles). Top: concentration profiles corresponding to nonnegativity constraints and nonnegativity plus unimodality constraints, bottom: spectral profiles corresponding to nonnegativity and nonnegativity plus unimodality constraints. All profiles are computed from the results presented in Fig. 6. In order to make the gap in the bands of profiles clear to the reader no additional lines are plotted as in Fig. 2. The blue (bottom, 2nd from right) resp. red (top and bottom, right) lines mark the band boundaries that are not computed by MCR-BANDS.

be resolved by the SCF approach since only two extrema are computed for each profile. The associated profiles are presented in Fig. 7.

#### 4.2. A three-component model problem with a split subset

The AFS-sets and the SCF-extrema are computed for the three-component data set as shown in Fig. 5 (2nd row). The model data consists of  $k = 201$  spectra each to  $n = 221$  wavelengths. In Fig. 8 the areas of  $\mathcal{F}_C$  for which the transformed concentration profiles are unimodal respectively not unimodal are plotted. The SCF-extrema and the AFS-sets are computed for nonnegativity constraints and for nonnegativity and unimodality constraints. Unimodality constraints are only applied for  $C$ . The results are shown in Fig. 9. Once again, the SCF-extrema and the AFS-result do not disagree: The SCF-extrema that are outside of the unimodality-AFS shift into the unimodality-AFS if unimodality constraints are applied. But nevertheless the SCF-extrema cannot resolve the separation effect since only two extrema are computed for each profile. The associated profiles are not plotted.

#### 4.3. Indirect separation of subsets of the spectral AFS

For both model problems the separation of a one of the subsets of the concentration factor AFS is caused by the direct restriction of the outer polygon in the concentration space, see e.g. Fig. 8. Further, for both model problems the direct separation of one subset of the concentration factor AFS results in an indirect separation of one subset of the spectral AFS. This indirect effect is caused by the cross-relationship due to duality [22, 27, 32]. For the three-component model problem the blue AFS subset splits up, see the right plot in Fig. 9. In these cases, MCR-BANDS don't work correctly and even may show unique solution incorrectly. The concentration profiles and pure component spectra obtained by the AFS-sets and MCR-BANDS under nonnegativity and unimodality constraints are presented in the Figures 10 and 11.

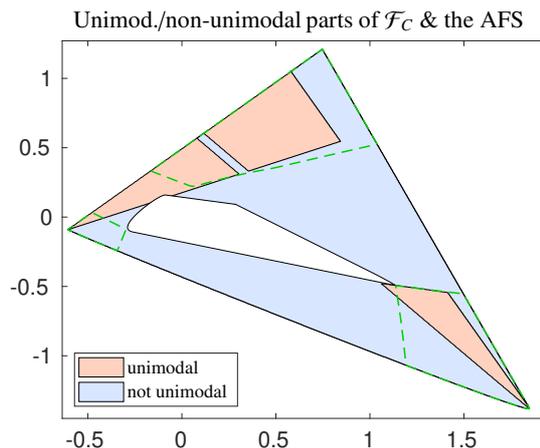


Figure 8: Unimodality and non-unimodality areas of outer polygon in the concentration space  $\mathcal{F}_C$  and their direct restrictions for the three-component model problem from Sec. 4.2. Points in the red colored area represent unimodal profiles and points in the blue colored area represent non-unimodal profiles. The green dashed lines are the boundaries of the AFS-sets for nonnegativity constraints. The inner polygon  $\mathcal{I}_C$  is the central white area.

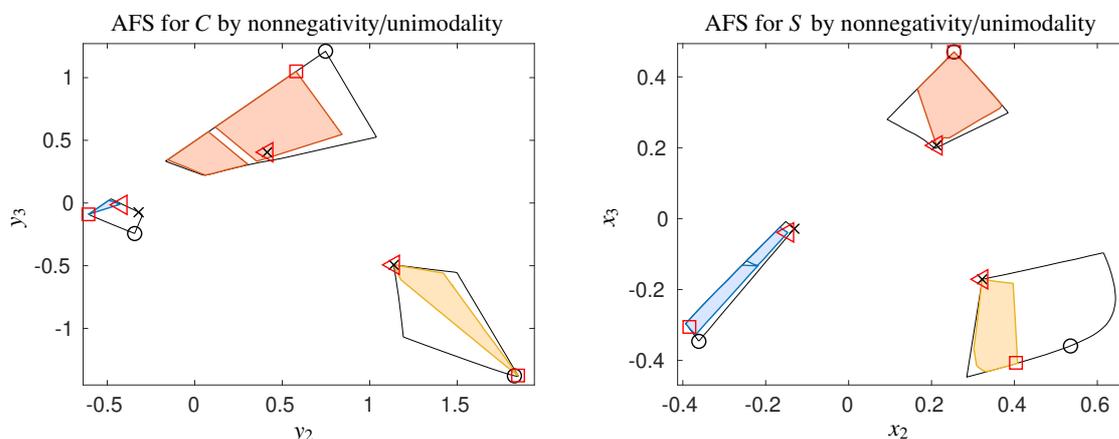
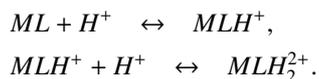


Figure 9: The AFS-sets and the SCF-extrema with (colored areas respectively red markers) and without (black lines respectively black markers) additional unimodality constraints for the data of Sec. 4.2. The SCF-extrema and the AFS-sets are not in mutual conflict. In order to compute also the separated subsets the polygon inflation algorithm was slightly modified.

## 5. Application to experimental data

The spectrophotometric titration of a solution containing  $4.56 \times 10^{-3}\text{M}$  ligand (3,7-diazanonanedioic acid di-amide),  $4.1 \times 10^{-3}\text{M}$   $\text{Cu}^{2+}$  and  $4.3 \times 10^{-3}\text{M}$   $\text{HCl}$  in the fixed ionic strength (0.5M  $\text{KCl}$ ) was carried out with 0.196M  $\text{NaOH}$ . During this titration, the pH of the solution was changed from 6.2 to 11.5 using  $k = 16$  values. The UV/Vis spectra were recorded between 450 nm and 750 nm using an equidistant frequency grid with  $n = 21$  points. The underlying kinetic reads



Therein  $ML$  is the neutral complex of  $\text{Cu(II)}$  which converts into its mono- and di-(carbonyl)-amide complexes using protonation. A hard modeling results in  $\log(K_1) = 8.5$  and  $\log(K_2) = 7.5$ . See [11] for more details on the data set. The mixed data as well as the profiles of the true factorization are presented in Fig. 12.

For these data not only unimodal concentration profiles can be expected but also unimodal pure component spectra. To compare the effect of unimodality constraints for both approaches (AFS respectively SCF-extrema) we use four different combinations of restrictions. The AFS-sets and SCF-extrema are computed by applying

1. nonnegativity constraints for both factors, see also the results from [11],
2. nonnegativity constraints for both factors and unimodality constraints for  $C$ .

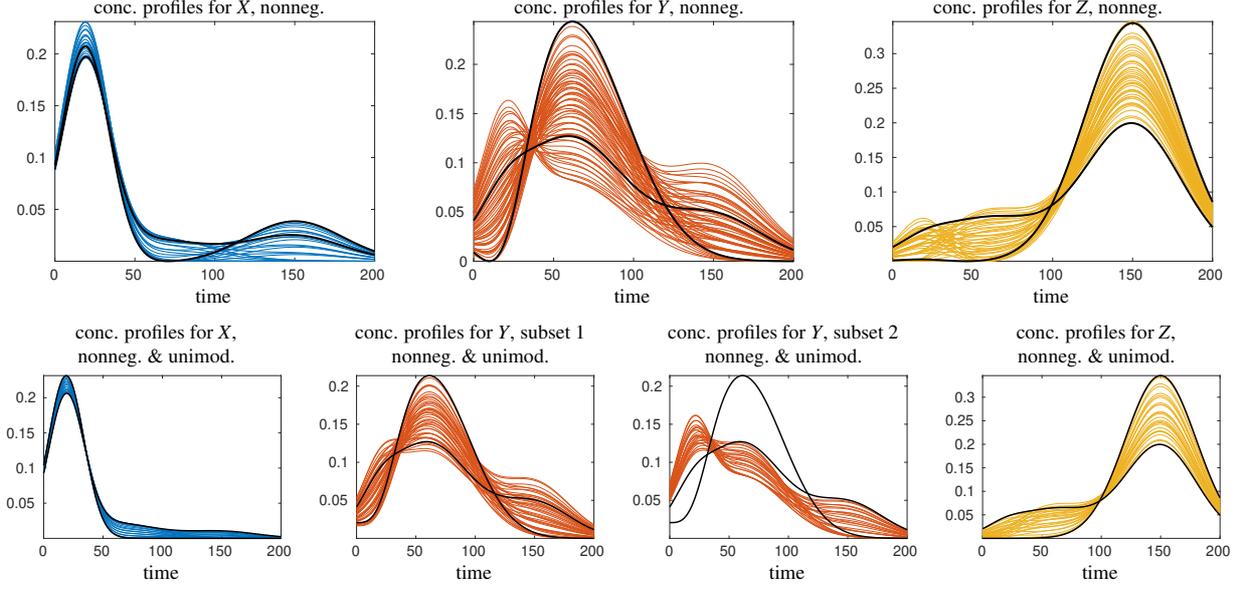


Figure 10: Feasible profiles for factor  $C$  (colored profiles) and MCR-BANDS extremal profiles (black profiles) corresponding to the single profiles ( $X$ ,  $Y$  and  $Z$ ). Top: profiles corresponding to nonnegativity constraints, bottom: profiles corresponding to nonnegativity and unimodality constraints – the profiles to component  $Y$  are presented in two single plots since the AFS segment splits up under additional unimodality constraints. All profiles are computed from the results of the left plot of Fig. 9.

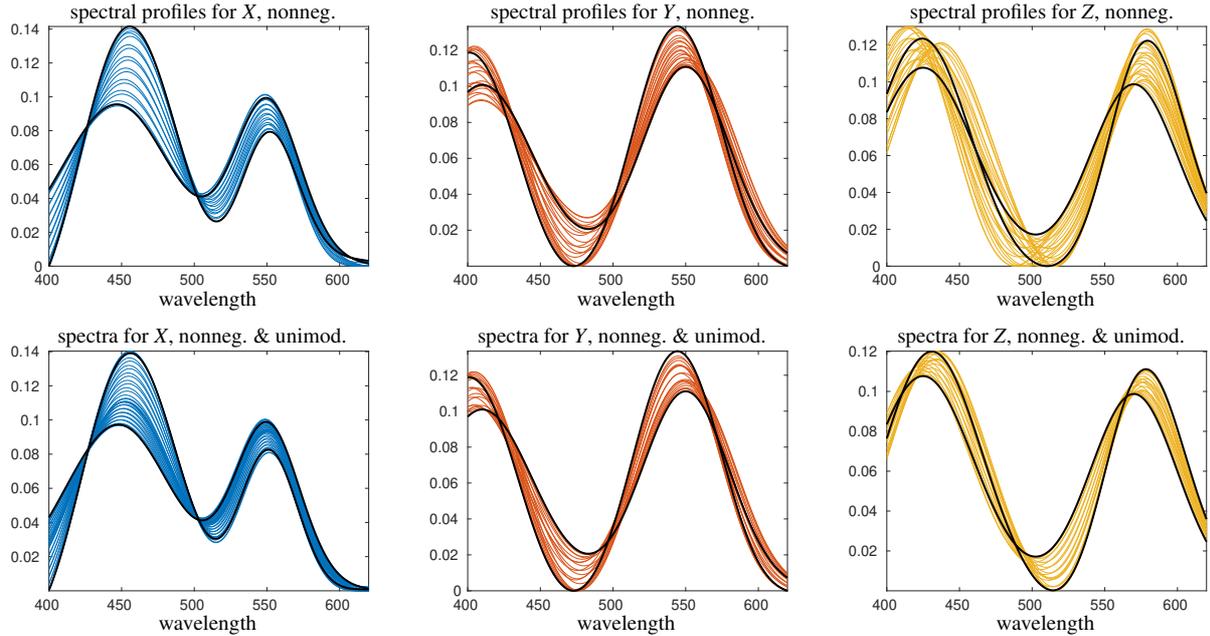


Figure 11: Feasible profiles for factor  $S^T$  (colored profiles) and MCR-BANDS extremal profiles (black lines) corresponding to the single components ( $X$ ,  $Y$  and  $Z$ ). Top: feasible spectra corresponding to nonnegativity constraints. Bottom: feasible spectra corresponding to nonnegativity and unimodality constraints. All profiles are computed from the results of the right plot of Fig. 9.

3. nonnegativity constraints for both factors and unimodality constraints for  $S$ .
4. nonnegativity and unimodality constraints for both factors. (The current MCR-BANDS implementation does not offer this opportunity. We modified the code with respect to the application of unimodality constraints in both directions.)

The restrictions of the outer polygons  $\mathcal{F}_C$  and  $\mathcal{F}_S$  are presented in Fig. 13. Since unimodal profiles can be expected for both factors the two outer polygons are separated in their areas which result in unimodal respectively non-unimodal profiles.

The AFS-sets as well as the SCF-extrema are computed for the four types of restrictions. The results are presented

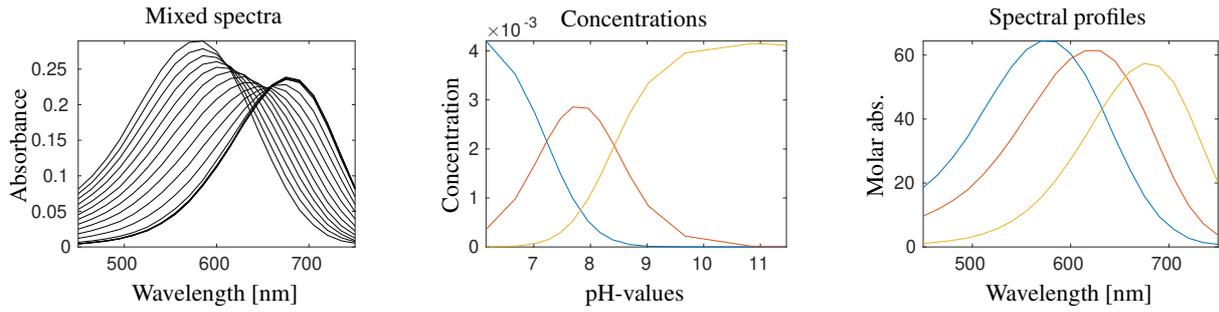


Figure 12: The mixed spectra and the true pure component profiles of the experimental data as introduced in Section 5.

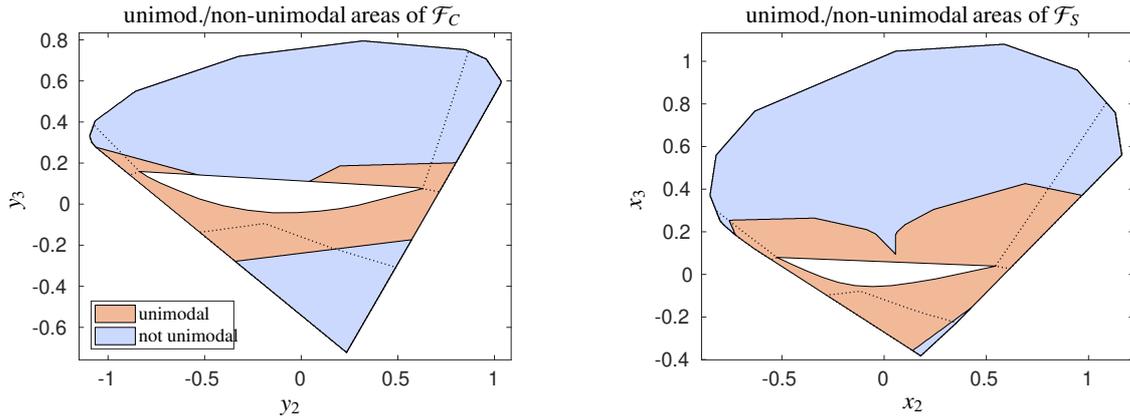


Figure 13: Unimodal (red) and non-unimodal (blue) areas of the outer polygons. The inner polygons are plotted white. The AFS-sets for applying only nonnegativity constraints are plotted as dotted black lines. Note that the restriction of both outer polygons by unimodality constraints is a special situation for this experimental data.

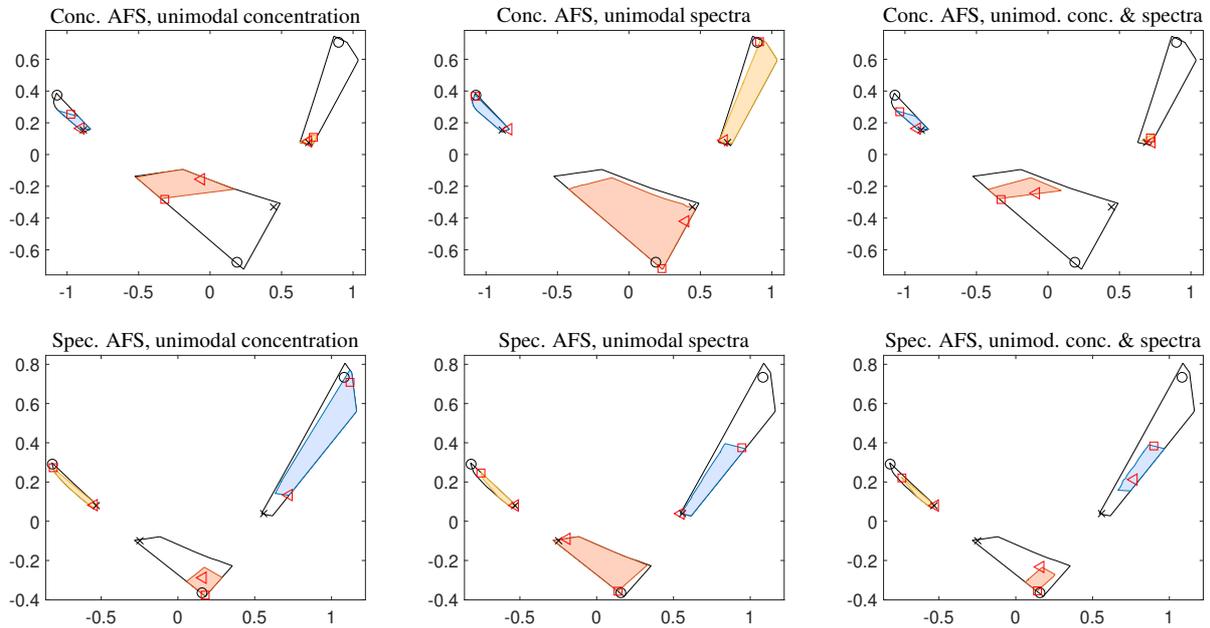


Figure 14: The AFS-sets and the representations of the SCF-extrema for the experimental data from Sec. 5 and different types of constraints. Top: concentration factor AFS-sets, 2nd row: spectral AFS-sets. Applied constraints next to nonnegativity for  $C$  and  $S$ : unimodality for  $C$  (left), unimodality for  $S$  (center), unimodality for  $C$  and  $S$  (right). The boundaries of the AFS-sets for applying only nonnegativity constraints are plotted by black lines and the restricted AFS-sets are plotted in color. The representations of the SCF-minimia (circles & squares) and SCF-maxima (crosses & triangles) are marked in black for applying only nonnegativity constraints respectively in red for applying additional unimodality constraints.

in Fig. 14. The AFS-sets and the representations of the SCF-extrema for applying only nonnegativity constraints are

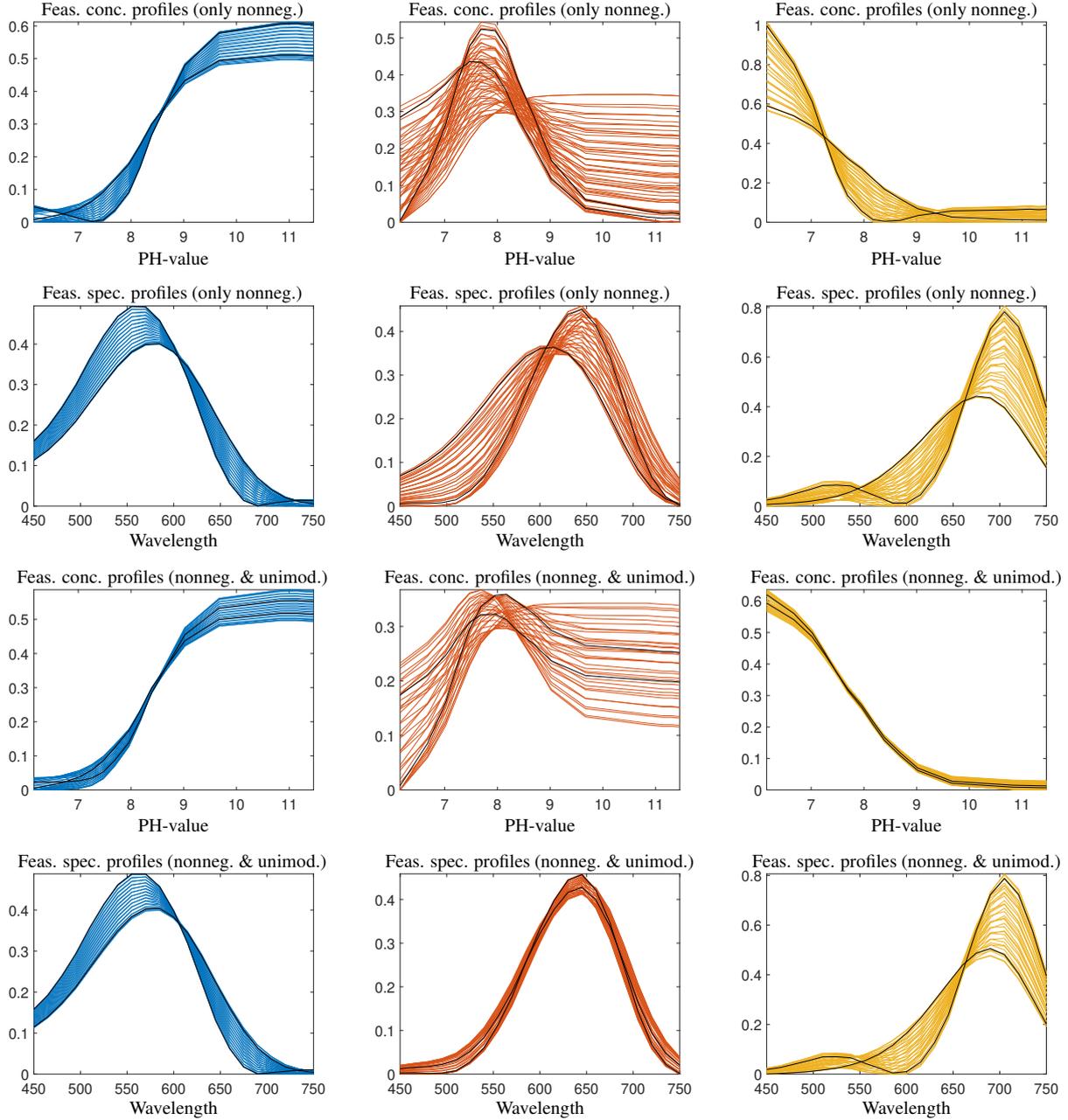


Figure 15: Feasible profiles for the experimental data from Sec. 5 using only nonnegativity constraints for both factors (1st & 2nd row) respectively additional unimodality constraints for both factors (3rd & 4th row). The extremal profiles from the SCF are plotted in black. The unimodality constraints exclude irrelevant profiles for both factors. All profiles are computed from the results of the right plots from Fig. 14.

plotted always in black. The results for the constrained cases are presented as colored areas (AFS-sets) and by red markers (for the representations of SCF-extrema). We see for this experimental data once again the reducing effect of the different types of unimodality constraints not only for the AFS-sets but also for the representations of the SCF-extrema. Once again, the SCF-extrema that do not represent unimodal profiles shift to the boundaries of the restricted AFS-sets if additional unimodality constraints are applied. The SCF can only resolve the ambiguity in form of the two extremal profiles. A problem for the MCR-BANDS computations under unimodality constraints is their dependency of the optimization on the initial guesses. Several computations were necessary with different initial factorizations. The feasible profiles and SCF-extrema for applying only nonnegativity constraints as well as for applying additionally unimodality constraints for  $C$  and  $S$  are presented in Fig. 15.

## 6. Conclusion

The problem of rotational ambiguities in the analysis of spectral data is implicit to all MCR methods. Polygon inflation and grid search are two methods for the exploration of the range of feasible solutions in terms of the AFS. The analysis of the SCF and the computation of its extrema, implemented in MCR-BANDS, focuses on the magnitude of ambiguity.

In this work, we compared the change of the AFS-sets and the SCF-extrema if unimodality constraints are applied next to nonnegativity constraints. It can be concluded that the representations of the SCF-extrema that are not part of the AFS-sets under unimodality constraints shift into these sets by applying unimodality constraints. However, the SCF-extrema do not result in the complete calculation of the AFS-sets and can only be considered as indicators of the rotational ambiguity. The effect of a separation of AFS-subsets, caused by unimodality constraints is presented and the results of the various approaches are discussed. A challenge for the further work is a modified geometric AFS construction by Borgen plots for such cases where AFS-subsets split up due to unimodality constraints. Another interesting question is how SCF-extrema behave in comparison to the AFS if monotonicity constraints are applied or the SCF is considered with respect to other norms.

## 7. Acknowledgement

This report was written during a research study of Somaye Vali Zade at the University of Rostock, Germany.

## 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. *Anal. Chem.*, 81(6):2115–2122, 2009.
- [2] H. Abdollahi and S. M. Sajjadi. Soft-modeling based spectrofluorimetric study of simultaneous equilibria. *Luminescence*, 24(5):332–339, 2009.
- [3] H. Abdollahi and R. Tauler. Uniqueness and rotation ambiguities in multivariate curve resolution methods. *Chemom. Intell. Lab. Syst.*, 108(2):100–111, 2011.
- [4] M. Akbari Lakeh, R. Rajkó, and H. Abdollahi. Local Rank Deficiency Caused Problems in Analyzing Chemical Data. *Anal. Chem.*, 89(4):2259–2266, 2017.
- [5] H. P. Bailey, S. C. Rutan, and P. W. Carr. Factors that affect quantification of diode array data in comprehensive two-dimensional liquid chromatography using chemometric data analysis. *J. Chromatogr. A*, 1218(46):8411–8422, 2011.
- [6] S. Beyramysoltan, H. Abdollahi, and R. Rajkó. Newer developments on self-modeling curve resolution implementing equality and unimodality constraints. *Anal. Chim. Acta*, 827:1–14, 2014.
- [7] O. S. Borgen and B. R. Kowalski. An extension of the multivariate component-resolution method to three components. *Anal. Chim. Acta*, 174:1–26, 1985.
- [8] S. D. Brown, R. Tauler, and B. Walczak. *Comprehensive chemometrics*. Elsevier Science, 2009.
- [9] A. De Juan, E. Casassas, and R. Tauler. Soft modeling of analytical data. *Encyclopedia of analytical chemistry: Applications, theory and instrumentation*, 2006.
- [10] Paul J Gemperline. Computation of the range of feasible solutions in self-modeling curve resolution algorithms. *Anal. Chem.*, 71(23):5398–5404, 1999.
- [11] A. Golshan, H. Abdollahi, S. Beyramysoltan, M. Maeder, K. Neymeyr, R. Rajkó, M. Sawall, and R. Tauler. A review of recent methods for the determination of ranges of feasible solutions resulting from soft modelling analyses of multivariate data. *Anal. Chim. Acta*, 911:1–13, 2016.
- [12] A. Golshan, H. Abdollahi, and M. Maeder. Resolution of rotational ambiguity for three-component systems. *Anal. Chem.*, 83(3):836–841, 2011.
- [13] G. H. Golub and C. F. Van Loan. *Matrix computations*, volume 3. JHU Press, 2012.
- [14] J. Jaumot and R. Tauler. MCR-BANDS: A user friendly MATLAB program for the evaluation of rotation ambiguities in Multivariate Curve Resolution. *Chemom. Intell. Lab. Syst.*, 103(2):96–107, 2010.
- [15] A. Jürß, M. Sawall, and K. Neymeyr. On generalized Borgen plots. I: From convex to affine combinations and applications to spectral dataSpectra. *J. Chemom.*, 29(7):420–433, 2015.
- [16] W. H. Lawton and E. A. Sylvestre. Self modeling curve resolution. *Technometrics*, 13(3):617–633, 1971.
- [17] M. Maeder and Y. M. Neuhold. *Practical data analysis in chemistry*, volume 26. Elsevier, 2007.
- [18] E. R. Malinowski. *Factor analysis in chemistry*. Wiley, 2002.
- [19] K. Neymeyr, A. Golshan, K. Engel, R. Tauler, and M. Sawall. Does the signal contribution function attain its extrema on the boundary of the area of feasible solutions? *Chemom. Intell. Lab. Syst.*, page 103887, 2019.
- [20] H. Parastar. Multivariate curve resolution methods for qualitative and quantitative analysis in analytical chemistry. In *Data handling in science and technology*, volume 29, pages 293–345. Elsevier, 2015.
- [21] M. N. Rahimdoost, M. Sawall, K. Neymeyr, and H. Abdollahi. Investigating the effect of flexible constraints on the accuracy of self-modeling curve resolution methods in the presence of perturbations. *J. Chemom.*, 30(5):252–267, 2016.
- [22] R. Rajkó. Natural duality in minimal constrained self modeling curve resolution. *J. Chemom.*, 20(3-4):164–169, 2006.
- [23] R. Rajkó. Computation of the range (band boundaries) of feasible solutions and measure of the rotational ambiguity in self-modeling/multivariate curve resolution. *Anal. Chim. Acta*, 645(1-2):18–24, 2009.
- [24] R. Rajkó. Studies on the adaptability of different Borgen norms applied in self-modeling curve resolution (SMCR) method. *J. Chemom.*, 23(6):265–274, 2009.

- [25] R. Rajkó and K. István. Analytical solution for determining feasible regions of self-modeling curve resolution (SMCR) method based on computational geometry. *J. Chemom.*, 19(8):448–463, 2005.
- [26] C. Ruckebusch and L. Blanchet. Multivariate curve resolution: a review of advanced and tailored applications and challenges. *Anal. Chim. Acta*, 765:28–36, 2013.
- [27] M. Sawall, A. Jürß, H. Schröder, and K. Neymeyr. On the analysis and computation of the area of feasible solutions for two-, three-, and four-component systems. In *Data Handling in Science and Technology*, volume 30, pages 135–184. Elsevier, 2016.
- [28] M. Sawall, A. Jürß, H. Schröder, and K. Neymeyr. Simultaneous construction of dual Borgen plots. I: The case of noise-free data. *J. Chemom.*, 31(12):e2954, 2017.
- [29] M. Sawall, C. Kubis, D. Selent, A. Börner, and K. Neymeyr. A fast polygon inflation algorithm to compute the area of feasible solutions for three-component systems. I: concepts and applications. *J. Chemom.*, 27(5):106–116, 2013.
- [30] M. Sawall, A. Moog, C. Kubis, H. Schröder, D. Selent, R. Franke, A. Brächer, A. Börner, and K. Neymeyr. Simultaneous construction of dual Borgen plots. II: Algorithmic enhancement for applications to noisy spectral data. *J. Chemom.*, page e3012, 2017.
- [31] M. Sawall and K. Neymeyr. A fast polygon inflation algorithm to compute the area of feasible solutions for three-component systems. II: Theoretical foundation, inverse polygon inflation, and FAC-PACK implementation. *J. Chemom.*, 28(8):633–644, 2014.
- [32] M. Sawall and K. Neymeyr. On the area of feasible solutions and its reduction by the complementarity theorem. *Anal. Chim. Acta*, 828:17–26, 2014.
- [33] M. Sawall and K. Neymeyr. A ray casting method for the computation of the area of feasible solutions for multicomponent systems: Theory, applications and FACPACK-implementation. *Anal. Chim. Acta*, 960:40–52, 2017.
- [34] M. Sawall, N. Rahimdoust, C. Kubis, H. Schröder, D. Selent, D. Hess, H. Abdollahi, R. Franke, A. Boerner, and K. Neymeyr. Soft constraints for reducing the intrinsic rotational ambiguity of the area of feasible solutions. *Chemom. Intell. Lab. Syst.*, 149:140–150, 2015.
- [35] M. Sawall, N.I. Rahimdoust, Ch. Kubis, H. Schroeder, D. Selent, D.r Hess, H. Abdollahi, R. Franke, A. Boerner, and K. Neymeyr. Soft constraints for reducing the intrinsic rotational ambiguity of the area of feasible solutions. *Chemom. Intell. Lab. Syst.*, 149:140–150, 2015.
- [36] A. N. Skvortsov. Estimation of rotation ambiguity in multivariate curve resolution with charged particle swarm optimization (cPSO-MCR). *J. Chemom.*, 28(10):727–739, 2014.
- [37] R. Tauler. Multivariate curve resolution applied to second order data. *Chemom. Intell. Lab. Syst.*, 30(1):133–146, 1995.
- [38] R. Tauler. Calculation of maximum and minimum band boundaries of feasible solutions for species profiles obtained by multivariate curve resolution. *J. Chemom.*, 15(8):627–646, 2001.
- [39] R. Tauler, S. Lacorte, and D. Barceló. Application of multivariate self-modeling curve resolution to the quantitation of trace levels of organophosphorus pesticides in natural waters from interlaboratory studies. *J. Chromatogr. A*, 730(1-2):177–183, 1996.
- [40] S. Vali Zade, K. Sawall, M. and Neymeyr, and H. Abdollahi. Introducing the monotonicity constraint as an effective chemistry-based condition in self-modeling curve resolution. *Chemom. Intell. Lab. Syst.*, 2018 submitted.
- [41] M. Vosough, C. Mason, R. Tauler, M. Jalali-Heravi, and M. Maeder. On rotational ambiguity in model-free analyses of multivariate data. *J. Chemom.*, 20(6-7):302–310, 2006.