On the analysis and computation of the area of feasible solutions for two-, three- and four-component systems.

Mathias Sawall[a], Annekathrin Jürß[a], Henning Schröder[a], Klaus Neymeyr[a,b]

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

Abstract

The area of feasible solutions (AFS) is a low-dimensional representation of all possible concentration factors or spectral factors in nonnegative factorizations of a given spectral data matrix. The AFS analysis is a powerful methodology for the exploration of the rotational ambiguity inherent to the multivariate curve resolution problem. Up to now the AFS has been studied for two-, three- and four-component systems:

1. The AFS for two-component systems was introduced by Lawton and Sylvestre in 1971. For these two-dimensional problems the AFS can be constructed analytically.
2. For three-component systems the AFS can either be constructed geometrically (classical approach by Borgen and Kowalski from 1985) or it can be computed by numerical algorithms. Various computational techniques have been suggested by different groups in the recent past.
3. For four-component systems a first numerical method for its computation has been published recently. A new polyhedron inflation algorithm is under development.

In this review paper we explain the underlying concepts of the AFS theory and its contribution to a deepened understanding of the multivariate curve resolution problem. A survey is given on various methods for the computation of the AFS for two-, three- and four-component systems. The focus is on methods which approximate the boundary of the AFS for three-component systems by inflating polygons and for four-component systems by inflating polyhedrons. Several numerical examples are discussed and the MatLab-toolbox FACPACK for these AFS-computations is presented.

Key words: spectral recovery, factor analysis, nonnegative matrix factorization, area of feasible solutions, Generalized Borgen Plot, complementarity and coupling.

1. Introduction

Multivariate curve resolution techniques serve to extract the pure component information from multivariate (spectroscopic) data. Typically, the data is taken by spectral observation of a chemical reaction system on a time × frequency grid. If k spectra are measured, each at n frequencies, then the resulting matrix D is a $k \times n$ matrix. The measured data result from a superposition of the pure component spectra. Multivariate curve resolution methods can be applied in order to extract the pure component spectra and the concentration profiles. The basic bilinear model underlying these methods is the Lambert-Beer law. In matrix notation, the Lambert-Beer law is a relation between D and the matrix factors $C \in \mathbb{R}^{k \times s}$ and $A \in \mathbb{R}^{s \times n}$ in the form

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

Therein s is the number of the pure or at least independent components. An error matrix $E \in \mathbb{R}^{k \times n}$ with entries close to zero can be added on the right-hand side of (1) in order to allow approximate factorizations in case of perturbed or noisy data D. In general, the matrices C and A are called abstract factors. One is interested in finding a factorization $D = CA$ with chemically interpretable C and A. Then the columns of $C \in \mathbb{R}^{k \times s}$ are the concentration profiles along the time axis of the pure components. And the rows of $A \in \mathbb{R}^{s \times n}$ are the associated pure component spectra.

The aim of a multivariate curve resolution (MCR) method is to determine the number s together with the

pure component factors *C* and *A*. Sometimes no additional information on the pure components is available. Then the MCR method only uses *D* for the pure component decomposition within a model-free approach. The main hurdle for any MCR technique is the so-called rotational ambiguity of the solution. See, e.g., [27, 28, 2] for an introduction to the ambiguity problem. By applying additional hard- or soft-constraints to the pure component factorization problem, one can often determine a single solution by means of a regularized optimization problem. In case of proper constraints this solution can be the chemically correct one. A large number of successful MCR methods has been developed. Some examples are methods as MCR-ALS [19], RFA [29], SIMPLISMA [52], BTEM [7] and PCD [32]. Alternatively, one can give up the aim to determine only a single solution by solving a regularized optimization problem. Instead, one can follow the global approach of determining the full range of all nonnegative factorizations *D* = *CA* with nonnegative rank-*s* matrices *C* and *A*. Such continua of possible nonnegative matrix factors can graphically be presented either by drawing the bands of possible concentration profiles together with the bands of possible spectra [49] or by plotting these sets of feasible factors by a certain low-dimensional representation, the so-called *Area of Feasible Solutions* (AFS). The aim of this paper is to provide a systematic introduction to the AFS concept together with many references to the chemometric literature. Important properties of the AFS are presented together with computational techniques for its numerical approximation.
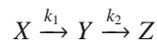
1.1. Organization of the paper

In the remaining part of this section, we introduce the four data sets which are used within this work. Section 2 explains in a compact form the principles of multivariate curve resolution methods and their relation to the singular value decomposition of the data matrix *D*. The techniques to compute the factors *C* and *A* are explained. In Section 3 the AFS and some of its important properties are described in detail. The rules for a classification of feasible or non-feasible points are discussed in Section 5. In Section 6 methods for the computation of the AFS are explained with a focus on the polygon inflation method. Methods for the reduction of the AFS by using additional information or soft constraints are demonstrated in Section 8. In Section 9 we study dynamic changes of the shape of the AFS under changes in the data (e.g. variation of a shift parameter). Finally in Section 10 the software package *FACPACK* for the computation of the AFS is pointed out.

1.2. Model data sets and experimental spectral data

In this work we use four different data sets (two model data sets and two IR-spectroscopic data sets) for all demonstrations. The data sets are as follows:

Data set 1 (FT-IR experimental data on a Rhodium catalyst formation)**.** *This data set describes the two-component subsystem (s = 2) of the formation of a Rhodium catalyst from a certain precursor; for the underlying chemical problem see [24]. The data set includes a number of k = 977 spectra, each with n = 481 wavenumbers. Within this spectral window the dominant absorbing components are a catalyst precursor and the catalyst. The usage of a first order reaction scheme allows to find unique factors C and A. These factors are shown in Figure 1. For the associated areas of feasible solutions see Figure 9.*

Data set 2 (A three-component model problem)**.** *The consecutive reaction*

$$X \xrightarrow{k_1} Y \xrightarrow{k_2} Z$$

is considered with the rate constants $k_1 = 0.3$ and $k_2 = 0.1$. The initial concentrations are $c_X(0) = 1$, $c_Y(0) = c_Z(0) = 0$ and the time interval is $t \in [0, 10]$. The solution of the associated rate equations and its discretization for a number of $k = 201$ time steps results in the concentration matrix factor $C \in \mathbb{R}^{201 \times 3}$. The matrix factor A is derived from the three Gaussian functions

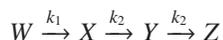$$a_X(x) = \exp\left(-\frac{(x-30)^2}{1250+w}\right), \quad a_Y(x) = \exp\left(-\frac{(x-50)^2}{1000+w}\right),$$

$$a_Z(x) = \exp\left(-\frac{(x-70)^2}{1000+w}\right)$$

for $x \in [0, 100]$ and equidistant discretization along the frequency axis. A number of $n = 401$ spectral channels is used. The spectra depend on the real parameter w which controls the signal width. The rows of the data matrix D as well as the original factors C and A for $w = 0$ are plotted in Figure 2. The associated areas of feasible solutions, namely the AFS for the concentration factors C and the AFS for the spectral factors A, are presented in the first column of Figures 20 and 21. See additionally the Figure 6 (with the data representation in the leftmost plot) and Figure 15 for a successive approximation of the AFS by means of the polygon inflation technique.

Data set 3 (Operando FT-IR specroscopic data from the Rhodium-catalyzed hydroformylation process)**.** *This data set is described in detail in [23]. The data consists*

of $k = 850$ *spectra, each with* $n = 642$ *wavenumbers. The three major absorbing components form a reaction-subsystem and are an olefin component, a hydrido-complex and an acyl-complex. All these chemical components are explained in [23]. The series of mixture spectra and the factors C and A are shown in Figure 3. The Michaelis-Menten kinetic has been used as a kinetic hard model in order to find unique concentration profiles. The complete set of all nonnegative factorizations of the data matrix D is represented by the AFS for the concentration factor and by the AFS for the spectral factor. These AFS sets are plotted together with the associated bands of feasible solutions in Figure 7.*

Data set 4 (A four-component model problem). *The concentration factor C of this model problem results from solving the rate equations for the reaction scheme*

$$W \xrightarrow{k_1} X \xrightarrow{k_2} Y \xrightarrow{k_2} Z$$

with the kinetic constants $k_1 = 1$, $k_2 = 0.25$ and $k_3 = 0.1$. The initial concentrations are $c_W(0) = 1$, $c_X(0) = c_Y(0) = C_Z(0) = 0$ and the time interval is $t \in [0, 10]$. The time-continuous concentration functions are discretized in $k = 26$ time steps to form C.

The factor A derives from the Gaussian functions

$$a_W(x) = \exp\left(-\frac{(x-40)^2}{\sigma}\right), \quad a_X(x) = \exp\left(-\frac{(x-20)^2}{\sigma}\right),$$

$$a_Y(x) = \exp\left(-\frac{(x-80)^2}{\sigma}\right), \quad a_Z(x) = \exp\left(-\frac{(x-60)^2}{\sigma}\right),$$

depending on the parameter σ on the frequency interval $x \in [0, 100]$. Numerical evaluation of these functions in $n = 31$ equidistant channels results in the matrix A. Thus the matrix A depends on σ. The mixture spectra, namely the rows of $D = CA$, are plotted together with the original factors C and A for $\sigma = 750$ in Figure 4. The associated areas of feasible solutions (concentrational and spectral AFS) are shown in Figure 19.

2. Multivariate curve resolution methods

Multivariate curve resolution methods are key-tools in order to extract the pure component information from the chemical mixture data in D. The problem is to compute

1. the number of independent components s and

2. the pure component factors C and A.

Any available information on the factors can and should be integrated into the MCR computations.

2.1. The singular value decomposition

The singular value decomposition (SVD), see [15], is a very powerful tool of numerical linear algebra to compute the left and right orthogonal bases for the expansion of the pure component factors $C \in \mathbb{R}^{k \times s}$ and $A \in \mathbb{R}^{s \times n}$; see for example [26, 28, 27, 38]. The SVD of D reads

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

Therein $U \in \mathbb{R}^{k \times k}$ and $V \in \mathbb{R}^{n \times n}$ are orthogonal matrices whose columns are the left and right singular vectors. The diagonal matrix Σ contains on its diagonal the singular values σ_i in decreasing order. The singular values are real and nonnegative. For an s-component system the first s singular vectors and the associated singular values contain all information on the system. For data not including perturbations only the first s singular values are nonzero if the chemical system contains s independent chemical components. For data including noise some additional singular values are nonzero. In such cases the SVD allows to compute optimal (with respect to least-squares) rank-s approximations of D. If in the case of noisy data the noise-to-signal ratio is not too large, then the number of independent chemical components s can often be determined from the SVD. Then the relevant and meaningful singular values are clearly larger compared to the remaining nonzero singular values which evince the influence of noise, cf. [27, 32].

2.2. Reconstruction of the pure component factors

The first s singular vectors, namely the first s columns of U and the first s columns of V, are used as bases to expand the desired pure component factors C and A. For ease of notation we denote these submatrices of the SVD-factors again by U and V. Then $U \in \mathbb{R}^{k \times s}$ and $V \in \mathbb{R}^{n \times s}$. The matrices C and A are formed according to

$$C = U\Sigma T^{-1}, \qquad A = TV^T. \tag{2}$$

Therein $T \in \mathbb{R}^{s \times s}$ is a regular matrix which remains to be determined. MCR methods typically provide a single pure component factorization $D = CA$ and thus they explicitly or implicitly determine the matrix T. From C and A the matrix T of expansion coefficients is accessible from Equation (2). For SVD-based MCR methods see [26, 28, 27, 32] and the references therein.

The basis expansion approach (2) drastically reduces the number of free variables of the pure component factorization problem. The crucial point is that the number of matrix elements of C and A is $(k + n)s$ whereas the
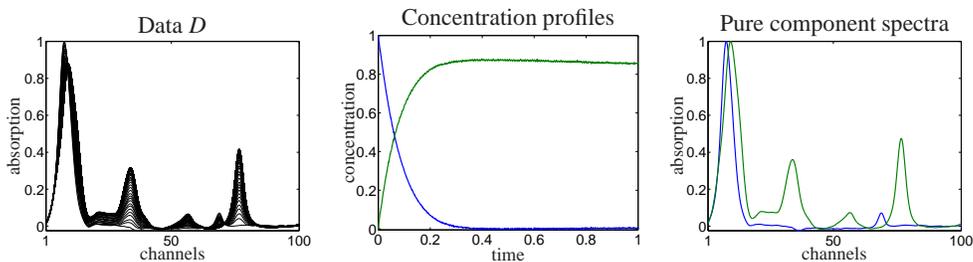
Figure 1: Data set 1: The series of mixture spectra is shown left (only every tenth spectrum of the data is plotted). All ordinate axes are scaled to a maximum of 1 and the channel windows are set to $[1, \ldots, 100]$. By enclosing a kinetic hard-model for the reaction scheme ($X \rightarrow Y$) into the pure component decomposition unique matrix factors (aside from the multiplicative ambiguity and the permutation ambiguity) have been determined. These are shown in the centered and right subplot.
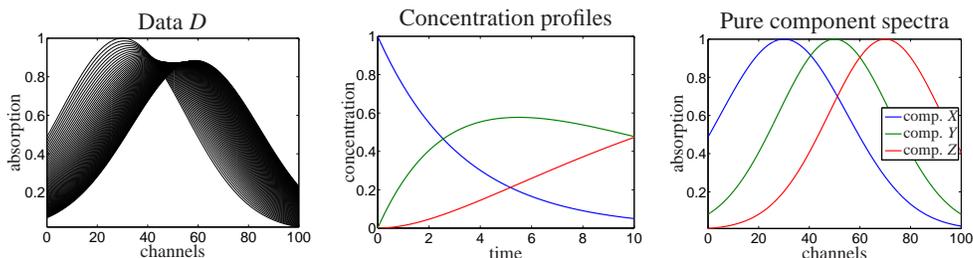

Figure 2: Data set 2 for $w = 0$: The leftmost plot shows the series of the mixture spectra, i.e. the rows of the matrix D. Only every third spectrum of the data is plotted. The remaining two plots show the concentration profiles and the spectra of the three pure components.
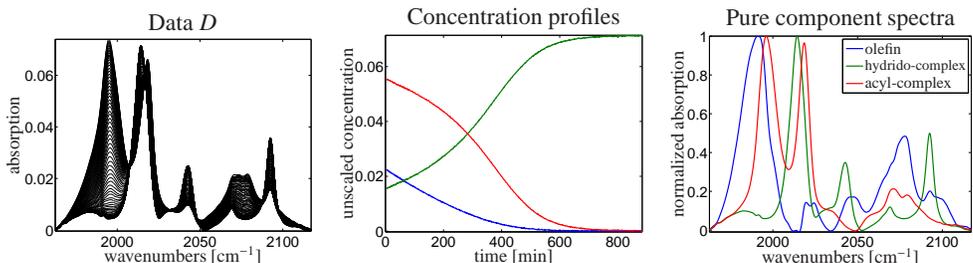

Figure 3: Data set 3 on catalyst formation within the hydroformylation process: The leftmost subplot shows the spectral data, i.e. the rows of D. Only every tenth spectrum of the data is actually plotted. The remaining two subplots show the concentration profiles and the spectra of the pure components. For a successful pure component decomposition of these spectral data a Michaelis-Menten kinetic hard-model has been integrated into the pure component factorization.
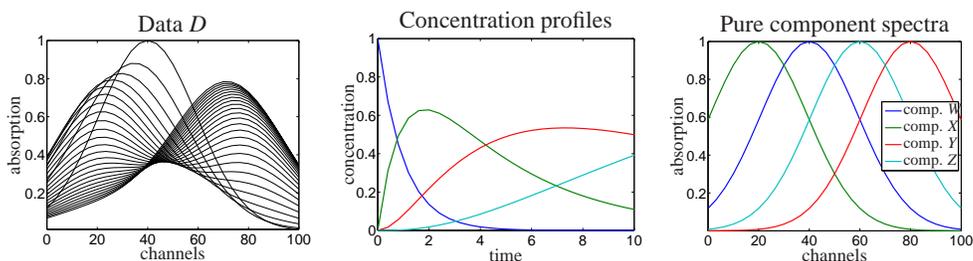

Figure 4: Data set 4 for $\sigma = 750$: The spectral data (left) for the four-component model problem together with the concentration profiles of the pure components (center) and the pure component spectra (right).

4

representation by Equation (2) reduces the degrees of freedom to the s^2 matrix elements of T. Hence the representation (2) is a basic ingredient for the construction of computationally effective MCR methods.

2.3. Application of hard- and soft constraints

Hard- and soft constraints have a crucial role in the construction of MCR methods [16, 11, 51, 32, 42]. A very restrictive and often successful hard constraint is a kinetic model of the underlying chemical reaction system. Only those concentration factors C are acceptable which are consistent with the kinetic model [8, 20, 27, 39]. Typically, the rate constants are implicitly computed as a by-product of the model fitting process.

If no kinetic model is available for C, then soft constraints can be used in order to extract (from the set of all nonnegative factorizations) solutions with special properties, see e.g. [16, 11, 51, 32, 42]. Typical examples of such soft constraints are those on the smoothness of the concentration profiles in C or A, constraints on a small or large integral of the spectra in A (in order to favor solution with few and sharp peaks or alternatively those with a large number of wide peaks), criteria on the closure of the concentration data and so on. Such soft constraints are usually added to the reconstruction functional $\|D - CA\|_F^2$ in terms of a cost function

$$g(T) = \sum_{i=1}^{p} \gamma_i \|g_i(C, A)\|_2^2.$$

According to the representations of $C = C(T)$ and $A = A(T)$ as functions of T, the cost function $g(T)$ includes a number of p constraint functions g_i. The $\gamma_i \geq 0$ are the associated weight factors which give the user the possibility to determine a certain balance between the different constraints. A small reconstruction error $\|D - CA\|_F^2$ together with the nonnegativity of the factors C and A is of highest importance; sometimes small negative entries in C and A can be acceptable. Other constraint functions are of lower importance, e.g. on the smoothness. For these constraint functions smaller weight factors γ_i are used.

3. The area of feasible solutions

Even with proper constraint functions and proper weight factors, MCR methods cannot always find the chemically correct or "true" solution. Thus one might follow the alternative idea to determine the set of *all* nonnegative factorizations $D = CA$. Such solutions which only fulfill the nonnegativity constraint are called *feasible* or *abstract* factors. The global approach of computing all feasible factors provides an elegant way in order to survey the complete rotational ambiguity of the pure component factorization problem. However, the sets of feasible matrices $C \in \mathbb{R}^{k \times s}$ or $A \in \mathbb{R}^{s \times n}$ are difficult to handle. The key idea to make these sets of feasible factors accessible is their low-dimensional representation in terms of the so-called *Area of Feasible Solutions* (AFS). The AFS refers to the representation of C and A as functions of T by Equation (2). Only a single row of T is sufficient to represent a feasible factor, see Section 3.2 for the details. In the following, our analysis aims at determining the AFS for the spectral factor A starting from a spectral data matrix D. This analysis can immediately be used to determine the AFS for the concentrational AFS containing the feasible factors C. Therefore we apply the procedure to the transposed data matrix D^T since in $D = CA$ the factors change their places by the transposition $D^T = A^T C^T$.

The aim of this section is to introduce the AFS and to discuss some of its important properties. We consider especially those properties which are decisive for an effective numerical computation of the AFS.

3.1. Development of the AFS concept and discussion of methods for its numerical computation

In this section a short overview is given on the development of the AFS concepts. These developments are closely related with the growth of effective numerical methods for its computation.

The AFS for two-component systems was first analyzed by Lawton and Sylvestre in 1971, see [26]. The Lawton-Sylvestre plot is a 2D plot of the set of the two expansion coefficients (with respect to the basis of singular vectors) which result in nonnegative matrix factors C and A. The Lawton-Sylvestre plot for a two component system consists of two cones whose boundaries can be computed analytically, see [50, 1, 35] and Section 4.

For three-component systems a direct analogue of the Lawton-Sylvestre plot would be a three-dimensional plot of feasible expansion coefficients. Such three-dimensional objects are somewhat more complicated to draw, to handle and to understand. However, there is a tricky dimension reduction (by a certain scaling) which allows to represent these AFS sets for three-component systems only by two expansion coefficients (and thus by plots in 2D). This was suggested by Borgen and Kowalski, who published in 1985 [6] a geometric construction of these 2D AFS plots for three-component systems. These plots are called Borgen plots. The mentioned dimension reduction can be explained by a typ-

ical Lawton-Sylvestre plot which is shown in Figure 5. For a two-component system the Lawton-Sylvestre plot consists of two cones. If the first expansion coefficient t_1 is fixed to 1, then the intersection of the dash-dotted line at $t_1 = 1$ with the two cones are two separated intervals. These two intervals are a 1D analogue of the Borgen plots. For a three-component system the 2D Borgen plot is the two-dimensional intersection of a plane at t_1 with a 3D generalization of a Lawton-Sylvestre plot.

The *geometric construction* of the Borgen plots is deepened by further concepts in [36] and [22]. In addition to the geometric constructions of the Borgen plots, various techniques for a numerical approximation of the AFS for ($s = 3$)-component systems have been devised. These are the grid search method (for two-component systems in [50, 1]), the triangle enclosure method [12] and the polygon inflation method [43, 44]. One benefit of the numerical methods compared to the geometric methods is that the numerical methods are able to compute the AFS in the presence of noise. Recently, the classical geometric construction has been generalized in a way which allows geometric AFS constructions for noisy data [22].

For systems with more than three components the problem of AFS computations is more complex and requires large computation times. For 4-component systems a generalized triangle enclosure method which uses a slicing of the 3D AFS has already been published [14]. The basic grid search algorithm and the polygon inflation method can also be extended to such higher dimensional problems.

AFS computations can be used to determine the so-called *feasible bands*. To this end, one can draw the continua of concentration profiles or the continua of possible spectra which are represented by the AFS. The resulting feasible bands visualize the rotational ambiguity. The MCR-Bands toolbox [49, 21] pursues the same objective, namely to construct the feasible bounds. However, the MCR-Bands approach does not require a previously constructed AFS. Instead, the band boundaries are constructed by a minimization respectively maximization of a properly constructed target function.

3.2. The set of feasible pure component spectra

Our focus is on the construction of the set of all possible pure component spectra. The set of feasible concentration profiles can be computed similarly; therefore the computational procedure is to be applied to D^T. The permutation of the columns of C and the same permutation applied to the rows of A does not provide any new information. This fact is known as the (trivial) permutation ambiguity. A consequence of this property is that

the problem to find all feasible factors A is equal to the problem to determine the set of all first rows of the feasible factors A. Hence the set of feasible pure component spectra (also called feasible bands) for an s-component system reads

$$\mathcal{A} = \{a \in \mathbb{R}^n : \text{ exist } C, A \geq 0 \text{ with } A(1, :) = a$$
$$\text{and } D = CA\}. \tag{3}$$

For the computation of $\mathcal{A}$ we prefer the SVD-based approach (2). A further reduction of the degrees of freedom is possible. This is explained next.

3.3. Reduction of the degrees of freedom

Equation (2) is a representation of the $s \times n$ matrix A by the matrix T with only s^2 matrix elements. These s^2 matrix elements are the expansion coefficients with respect to the basis of right singular vectors. As shown for the derivation of (3) only the first row of $A = TV^T$ is required in order to form the set $\mathcal{A}$ of feasible spectra. The first row of A equals the first row of TV^T. Hence only the s matrix elements of the first row of T are decisive. This reduces the degrees of freedom from s^2 down to s.

These remaining s degrees of freedom for an s-component reaction system can further be reduced to $s - 1$ by a certain scaling of the rows of A. In [6, 36] the $\| \cdot \|_1$ vector norm (i.e. the sum of the absolute values of the components) is used for the normalization of the spectra. Alternatively, the maximum norm $\| \cdot \|_{\max}$ can be used. Here we follow the approach in [12, 14, 43, 44] and use a scaling which sets the first column of T equal to ones, i.e.

$$T(i, 1) = 1 \qquad \text{for all } i = 1, \ldots, s. \tag{4}$$

This can be called the first-singular-vector scaling (FSV-scaling) since it uses for the first right singular vector the fixed expansion coefficient 1. A precise justification for this choice is based on the Perron-Frobenius theory [30] on spectral properties of nonnegative matrices; for the details see Theorem 2.2 in [44]. Figure 5 shows a typical Lawton-Sylvestre plot for a two-component model problem. The dash-dotted and the dotted lines define two intersections with the Lawton-Sylvestre plot. These intersections are the 1D AFS representations with respect to the FSV-scaling and with respect to the $\| \cdot \|_1$ normalization.
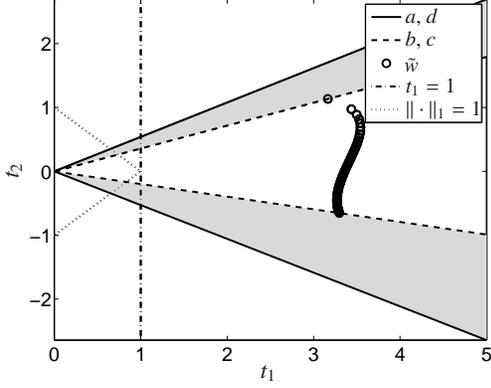
Figure 5: The Lawton-Sylvestre plot (gray triangle-shaped areas) for a two-component model problem. The dash-dotted and the dotted lines define two intersections with the Lawton-Sylvestre plot. These intersections are the 1D AFS representations with respect to the FSV-scaling and with respect to the $\| \cdot \|_1$ normalization. The points $\tilde{w}(i,:) = D(i,:)V(:,1:2)$, $i = 1, \ldots, n$, are drawn by small circles. These points represent the rows of D. Mathematically the coordinates of these points are the expansion coefficients of the rows of D with respect to the two dominant right singular vectors. For a more detailed explanation of the AFS and for the meaning of a, b, c and d see Section 4.

3.4. Definition of the AFS

With respect to the FSV-scaling (4) the matrix T in (2) has the form

$$
T = \begin{pmatrix} 1 & x_1 & \cdots & x_{s-1} \\ 1 & & & \\ \vdots & & S & \\ 1 & & & \end{pmatrix}.
\tag{5}
$$

Therein S is an $(s-1) \times (s-1)$ matrix. Only the $s-1$ elements of the row vector $x = (x_1, \ldots, x_{s-1})$ are decisive for the representation of the set of all feasible solutions.

With these definitions the set $\mathcal{A} \subset \mathbb{R}^n$ by Equation (3) can be represented by the associated set of expansion vectors $x \in \mathbb{R}^{s-1}$. Such a set of $(s-1)$-dimensional vectors for a chemical reaction system with s species is much easier to handle compared to the subset $\mathcal{A}$ of the higher dimensional space $\mathbb{R}^n$. The set

$$
\mathcal{M} = \{x \in \mathbb{R}^{s-1} : \text{ exists } S \text{ so that } T \text{ in (5) fulfills}
$$
$$
\text{rank}(T) = s \text{ and } C, A \geq 0\}
\tag{6}
$$

is called the AFS for the factor A or the spectral AFS.

Figure 6 shows a typical AFS for the three-component underlying the data set 2. This AFS consists of three isolated subsets. In [43, 44] these subsets are called segments of an AFS. Further, two series of points are marked within two segments of the AFS. The associated series of spectra, which are represented by these

points, are also shown in Figure 6. Figure 7 displays the AFS sets of the factors C and A for the data set 3. Additionally, this figure shows the associated bands of feasible pure component spectra, i.e. the set $\mathcal{A}$. The feasible bands of concentration profiles are also plotted; this set results from computing the set $\mathcal{A}$ for the transposed data matrix D^T.

3.5. Properties of the AFS

The AFS has several interesting properties. Many of these properties derive from the Perron-Frobenius spectral theory of nonnegative matrices [30]. This theory provides (see Section 3.3) the justification for the scaling condition (4). An important property of the AFS is its boundedness, see Section 3.5.2. This property makes possible a numerical approximation of the boundary of the AFS. The AFS sets may have several shapes. For three-component systems the most important cases are AFS sets which consist of three separated segments and AFS sets which have the form of a topologically connected set with a single hole. Such a hole always contains the origin (null vector), see, e.g., Figure 24 and Section 3.5.3. Further explanations on the geometric construction of the AFS and its relationship to the theory of simplices and convex combinations are contained in [36, 22].

3.5.1. Definition of FIRPOL and INNPOL

For the further analysis the two polygons FIRPOL and INNPOL are to be introduced. First, the set

$$
\mathcal{M}^+ = \{x \in \mathbb{R}^{s-1} : (1, x)V^T \geq 0\}
\tag{7}
$$

is called FIRPOL [6, 36, 22]. All points x in $\mathcal{M}^+$ result in nonnegative linear combinations of the right singular vectors, i.e. $(1, x)V^T$. Thus FIRPOL is a superset of the set of feasible spectra. The membership of a certain x to $\mathcal{M}^+$ does not guarantee that the nonnegative spectrum $(1, x)V^T$ is part of a feasible pure component decomposition $D = CA$. The crucial point is that nonnegativity of $(1, x)V^T$ does not necessarily imply the nonnegativity of an associated concentration profile.

Further, the set $\mathcal{M}^*$ is to be introduced

$$
\mathcal{M}^* = \{x \in \mathbb{R}^{s-1} : \text{ exists } S \text{ so that } T \text{ in (5) fulfills}
$$
$$
\text{rank}(T) = s \text{ and } C \geq 0, \ A(2:s,:) \geq 0\}.
\tag{8}
$$

The two sets $\mathcal{M}^+$ and $\mathcal{M}^*$ are super-sets of the AFS $\mathcal{M}$. The definition of $\mathcal{M}^+$ includes only the nonnegativity-constraint $A(1,:) \geq 0$. The definition of $\mathcal{M}^*$ includes
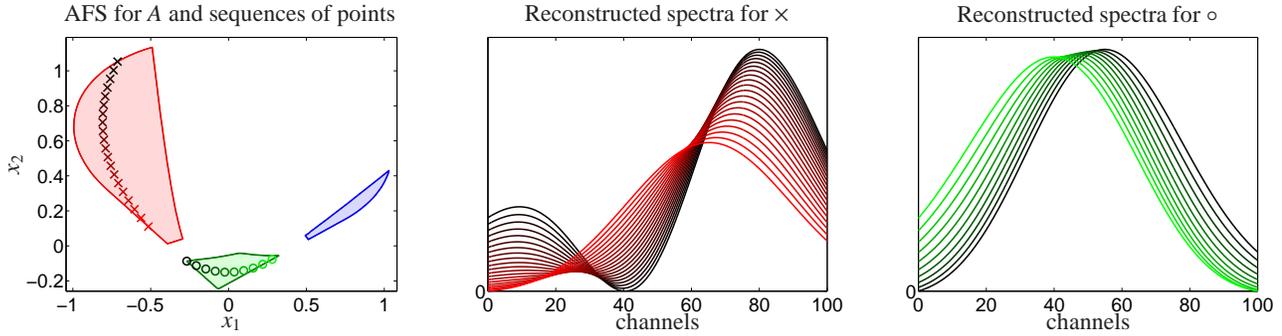
7

Figure 6: Data set 2 Left: The spectral AFS for this three-component system consists of three isolated subsets. In two of these subsets sequences of points are marked. The points marked by × have a color shading from red to black. The points marked by ○ have a color shading from green to black. The associated series of spectra are shown in the same color shading in the remaining two subplots.
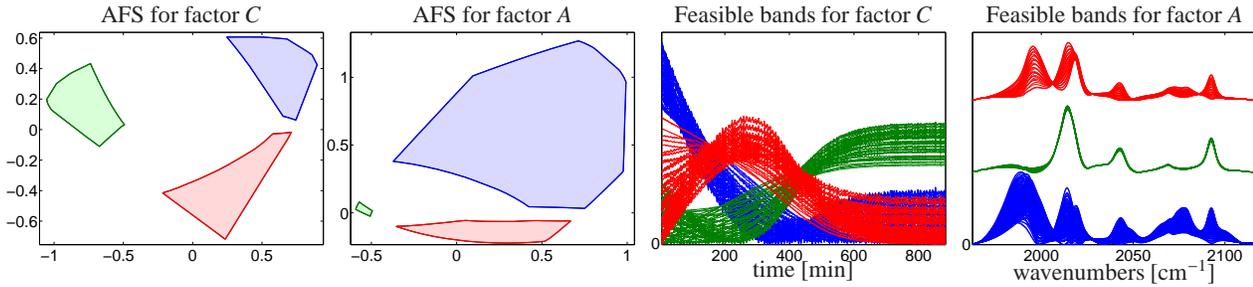

Figure 7: Data set 3: The areas of feasible solutions, namely the AFS for the concentration factor C and the AFS for the spectral factor A, are plotted (left) together with the associated bands of feasible solutions (right).

the remaining constraints on nonnegativity and the rank condition. Thus $\mathcal{M} = \mathcal{M}^+ \cap \mathcal{M}^*$. Finally, the vectors

$$w(i, :) = \frac{D(i, :)V(:, 2 : s)}{D(i, :)V(:, 1)} \in \mathbb{R}^{s-1} \text{ for } i = 1, \ldots, n \quad (9)$$

are introduced. In [6, 36, 22] the convex hull of these points $w(i, :)$, $i = 1, \ldots, n$, is called INNPOL [6, 36, 22].

3.5.2. Boundedness of the AFS

Three *numerical* approximation methods for the computation of the AFS for two- and three-component system have been described in literature. These are the grid search algorithm [50], the triangle enclosure method [12] and the polygon inflation method [43]. In order to guarantee that these algorithms terminate within finite times, the boundedness of the AFS is required. Theorem 2.4 in [44] proves that $\mathcal{M}^+$ is a bounded set if and only if $D^T D$ is an irreducible matrix. Thus the AFS $\mathcal{M}$ is also a bounded set since $\mathcal{M} \subset \mathcal{M}^+$. See [30] or [44] for the definition of an irreducible quadratic matrix. Practically, irreducibility is always guaranteed if the whole system does not allow a complete separation into two independent subsystems. The latter case is al-

ready a trivial case as completely isolated subsystems can be analyzed separately.

3.5.3. The origin is never included in the AFS

As already mentioned in Section 3.5 the origin (or null vector) is never contained in the AFS. The proof for this fact is given in Theorem 2.2 in [44]. It is based on the Perron-Frobenius theory and uses the irreducibility of the matrix $D^T D$. For the cases that the AFS consists of several isolated subsets, called segments, these subsets do not contain any holes. In mathematics such sets are called *simply-connected*. The approach of inflating polygons can be used in order to approximate such AFS segments. The remaining case that the AFS consists of only a single set is more complicated. Such a single-segment AFS always contains a hole and this hole encloses the origin. The inverse polygon inflation algorithm is a modification of the polygon inflation algorithm which allows a fast and accurate numerical approximation of such one-segment AFS sets with a hole. See Section 4 in [44] for the details.

3.5.4. Geometric AFS construction

The geometric construction of the AFS for three-component systems was introduced by Borgen and Kowalski in 1985 [6, 5], see also [36, 22]. The resulting geometric constructions are called Borgen plots.

The construction principles of Borgen plots are not limited to $s = 3$ but can be applied to general s-component systems. In the general case a point x is feasible if and only if there exist further $s - 1$ points y_j, $j = 1 \ldots, s - 1$, so that the simplex spanned up x and the vectors y_j is enclosed in $\mathcal{M}^+$ and includes all points $w(i, :)$ given by Equation (9). These points are the vertices of INNPOL. Consequently the classical theory by Borgen and Kowalski works with convex equations. A weakening from *convex combinations* towards *affine combinations* allows to generalize the geometric AFS construction [22]. The resulting generalized Borgen plots can be constructed even for noisy or perturbed spectral data.

3.6. Segment structure of the AFS

For two-component systems and with respect to the FSV-scaling, the AFS (6) always consists of two separated 1D intervals. The intervals may degenerate to single points. One of these intervals is completely negative and the other one is completely positive, see Figure 8 and in Figure 5 the cross-section at $t_1 = 1$.

For three-component systems the AFS can consist of a single segment with a hole around the origin or of a number of $3m$ segments for $m = 1, 2, \ldots$. A formal proof is planned for a forthcoming paper. For experimental data sets only the cases of a one-segment AFS and of 3-segment AFS sets have been observed. Only these cases appear to be practically relevant. However, nonnegative matrices D can be constructed whose associated AFS sets consist of 6, 9, $\ldots$ separated segments.

For s-component systems with $s \geq 4$ little information exists on the possible numbers of segments.

By changing the data D continuously one can explore the resulting changes of the associated AFS sets. However, such parameter dependent data sets could hardly be found as experimental data sets. The data sets 2 and 4 in Section 1.2 include the parameters w and σ. For each of these data sets a variation of these parameters allows to start from AFS sets with either 3 or 4 separated segments and to end in one-segment AFS sets each with a hole. See Figure 24 for the AFS-dynamics in case of the data set 2 and Figure 25 for the data set 4.

4. The AFS for two-component systems

The AFS for systems with a number of $s = 2$ components can explicitly be described analytically. The numerical evaluation of the analytic formula results in the 1D AFS plots. Next these formula are compiled. The starting point for the case $s = 2$ is the matrix T by (5) which together with its inverse read

$$T = \begin{pmatrix} 1 & x \\ 1 & S_{11} \end{pmatrix}, \qquad T^{-1} = \frac{1}{S_{11} - x} \begin{pmatrix} S_{11} & -x \\ -1 & 1 \end{pmatrix}.$$

The nonnegativity for the factors C and A results in feasible intervals for x and S_{11}, see also Section 3.6 in [44] and [1]. With

$$a = -\min_{i \text{ with } V(i,2)>0} \frac{V_{i1}}{V_{i2}}, \qquad d = -\max_{i \text{ with } V(i,2)<0} \frac{V_{i1}}{V_{i2}},$$
$$b = \min_i \frac{U_{i2}\sigma_2}{U_{i1}\sigma_1}, \qquad c = \max_i \frac{U_{i2}\sigma_2}{U_{i1}\sigma_1} \tag{10}$$

the AFS for the two-component system has the form of two separated intervals

$$\mathcal{M} = [a, b] \cup [c, d]. \tag{11}$$

This result can be interpreted in a way that both

$$x \in [a, b] \quad \text{and} \quad S_{11} \in [c, d]$$

and

$$S_{11} \in [a, b] \quad \text{and} \quad x \in [c, d]$$

result in nonnegative factors.

A certain choice

$$(x, S_{11}) \in [a, b] \times [c, d]$$

completely determines a nonnegative factorization $D = CA$. The second choice $(S_{11}, x) \in [a, b] \times [c, d]$ does not provide any new information. Instead, the second solution is equal to the first solution after a row permutation in A and a column permutation in C. This fact justifies that the AFS for two-component systems is often represented by the rectangular $[a, b] \times [c, d]$, see [50, 1, 44].

4.1. Numerical AFS computation for the data set 1

For noisy or perturbed spectral data it can be advantageous to accept small negative entries in the pure component factors. With a positive control parameter ε on the tolerated size of negative entries of C and A (see Section 5.1) one can generalize the AFS-bounds as discussed above. Next numerical results generated by the *FACPACK* software [45] are presented for the two-component experimental FT-IR spectroscopic data set
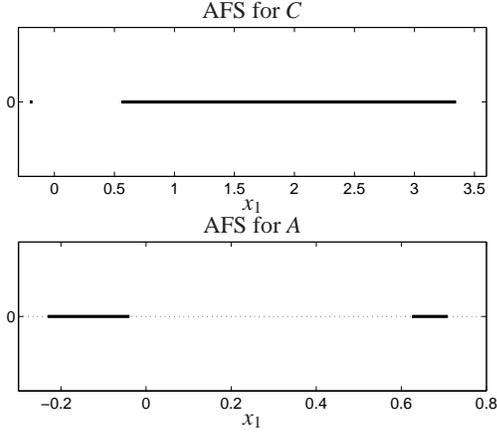
AFS for C

AFS for A

Figure 8: Data set 1: The AFS $\mathcal{M}$ for this two-component system for the concentrational factor (upper plot) and for the spectral factor (lower plot). The control parameter ε on acceptable negative entries reads $\varepsilon = 0.035$.

Rectangle-AFS for C
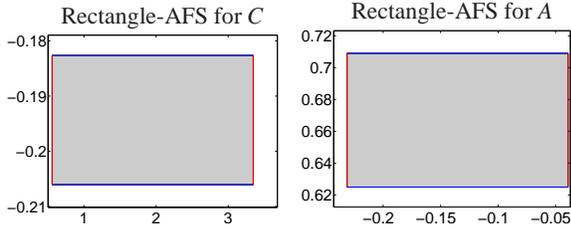
Rectangle-AFS for A

Figure 9: Data set 1: Two-dimensional rectangle representation of the 1D-AFS in which the two intervals of the one-dimensional AFS $\mathcal{M}$ are the edges of a rectangular. A certain point (with the two coordinates x and S_{11}) completely determines either C and A. The control parameter is again $\varepsilon = 0.035$.

1. Figure 8 shows the one-dimensional AFS sets with $\varepsilon = 0.035$ for the concentrational factor C and for the spectral factor A. Additionally, Figure 9 shows for the same problem the two-dimensional rectangle representation of the same AFS. The two intervals of the one-dimensional AFS form the edges of a rectangular. The advantage of such a rectangle-representation is that a certain point (with the two coordinates x and S_{11}) completely determines C or A. Moreover, a known factor C allows to compute A from $D = CA$ and vice versa.

5. Feasibility of points in the AFS

For the case of two-component systems the question of the feasibility of a certain point is solved by the analysis in Section 4 (at least for noise-free data). In the following we consider systems with three or more components. For these systems the feasibility question arises in two major ways. There is first the feasibility analysis

for noise-free (model) data by considering a certain geometric construction. Alternatively, and with a focus on experimental spectral data, there is a numerical feasibility analysis which is based on the numerical solution of an optimization problem. Unfortunately, the numerical feasibility test can yield false results, if the numerical optimization procedure (e.g. due to convergence problems or poor initial estimates) is not successful.

This section explains the feasibility checks of the polygon inflation algorithm by soft constraints (in the Subsection 5.1), of the triangle enclosure technique as well as the grid search method (see Subsection 5.2) and of the geometric-constructive Borgen plot approach (see Section 5.3).

5.1. Soft-constraint based feasibility check

Soft constraints can be added to the feasibility check on nonnegativity. The aim of this approach is to compute the matrix elements of the submatrix S of (5) by solving a minimization problem for a certain target function which guarantees that CA is a good approximation of the initial matrix D. Simultaneously various constraints on C and A are to be satisfied. This feasibility test which also underlies the polygon inflation algorithm [44] is explained in the following.

First we introduce a small control parameter $\varepsilon \geq 0$ so that $-\varepsilon$ is a lower bound for the acceptable negative elements of C and A (in a relative sense related to the maximal value of a concentration profile or spectrum). Mathematically, these conditions read

$$\frac{\min_j C(j, i)}{\max_j C(j, i)} \geq -\varepsilon, \qquad \frac{\min_j A(i, j)}{\max_j A(i, j)} \geq -\varepsilon$$

for all $i = 1, \ldots, s$. The acceptance of small negative entries can often significantly stabilize the computation in case of noisy or perturbed (e.g. by a background subtraction) data.

The feasibility test for a point x is done in two steps. First, a rapid and computationally very cheap test is used in order to check whether x is contained in the set FIRPOL $\mathcal{M}^+$, see Equation (7). If x is not in FIRPOL, then x cannot be an element of the AFS $\mathcal{M}$. Once again, we accept small negative entries. To this end we use an approximate FIRPOL test in order to check whether or not

$$f_0(x) := \min\left(\frac{(1, x)V^T}{\|(1, x)V^T\|_\infty} + \varepsilon, \ 0\right) \qquad (12)$$

satisfies that $f_0(x) \geq 0$ (in a component-wise sense). Therein, $\|\cdot\|_\infty$ is the maximum vector norm, which is the largest absolute value of all components of its argument.
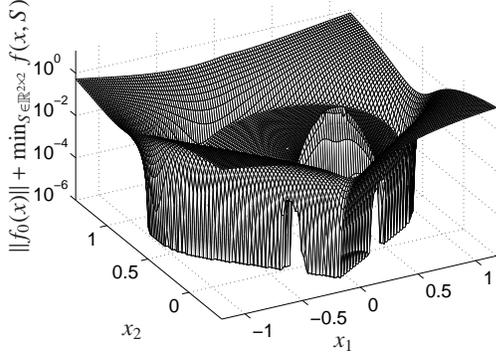
10

Figure 10: Data set 2: The function $\|f_0(x)\| + \min_S f(x, S)$ is plotted. The valley bottom is the AFS.

Figure 11: Data set 1: The function $\|f_0(x)\| + f(x, S)$ with scalar arguments x and S is plotted. The control parameter ε is set to 0.035 in order to successfully deal with small negative entries which are caused by a baseline correction. Since $f(x, S) = f(S, x)$ the function graph of $\|f_0\| + f$ is symmetric to the axis $x = S$; the symmetric part of the function is not plotted. The area with $\|f_0(x)\| + \min_S f(x, S) < \varepsilon_{\text{tol}}$ is bounded by the interval endpoints a, b, c and d as given in Equation (10).

If this test is passed successfully, then a second much more expensive test follows. Therefore the soft constraint function

$$
\begin{aligned}
f(x, S) = & \sum_{i=1}^{s} \left\| \min\left(\frac{C(:, i)}{\|C(:, i)\|_\infty} + \varepsilon, \ 0 \right) \right\|_F^2 \\
& + \sum_{i=2}^{s} \left\| \min\left(\frac{A(i, :)}{\|A(i, :)\|_\infty} + \varepsilon, \ 0 \right) \right\|_F^2 + \|I_s - TT^+\|_F^2
\end{aligned}
\tag{13}
$$

is considered with $T = T(x, S)$ by (5). Further, C, A are computed according to (2). If

$$
\min_{S \in \mathbb{R}^{(s-1) \times (s-1)}} f(x, S) \le \varepsilon_{\text{tol}},
\tag{14}
$$

then x has passed the feasibility test successfully. Therein ε_{tol} is a small positive control parameter, e.g. $\varepsilon_{\text{tol}} = 10^{-10}$.

To summarize, the approximate feasibility test with the control parameters ε and ε_{tol} results in the (approximate) AFS

$$
\begin{aligned}
\mathcal{M} = \{ x \in \mathbb{R}^{s-1} : \ & x \text{ fulfills } f_0(x) \ge 0 \\
& \text{and } \min_S f(x, S) \le \varepsilon_{\text{tol}} \}.
\end{aligned}
\tag{15}
$$

Figure 10 shows the function $\|f_0(x)\| + \min_S f(x, S)$ on the grid $(x_1, x_2) \in [-1.2, 1.2] \times [-0.4, 1.4]$ for the data set 2. For these computations the control parameters are $\varepsilon = 10^{-12}$ and $\varepsilon_{\text{tol}} = 10^{-6}$. All points with $\|f_0(x)\| + \min_S f(x, S) \le 10^{-6}$ belong to the AFS; these are the points at the valley bottom in Figure 10.

For the data set 1, a two-component system, Figure 11 displays the function $\|f_0(x)\| + f(x, S)$ on the grid $(x, S) \in [-0.29, -0.01] \times [0.55, 0.8]$ for $\varepsilon = 0.035$. For this two-component system no minimization is required; the arguments S and x of f are real numbers.
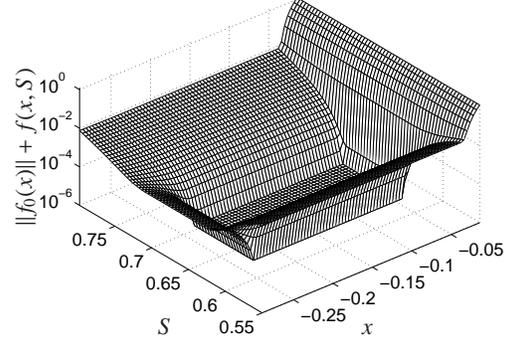
The valley bottom is equal to the right subplot of Figure 9. The endpoints of the intervals are approximately equal to a and b in (10) and are located on the x_1-axis. Further the endpoints c and d are located on the S-axis.

5.2. The ssq-function based feasibility check

The *ssq*-function (ssq for sum-of-squares) approach evaluates the reconstruction functional

$$
ssq(x, S) = \|D - \max(C, 0) \cdot \max(A, 0)\|_F^2.
$$

Therein $\max(C, 0)$ and $\max(A, 0)$ are the matrices whose negative entries are zeroed. The matrices C and A depend on $T = T(x, S)$ according to (5). The triangle enclosure algorithm [12] and the grid search method [50, 1] are based on the evaluation of the *ssq*-function. Computationally, the *ssq*-evaluation is relatively expensive as the computation of $O(k \cdot n)$ squares is required whereas the evaluation of (13) needs only $O(k + n)$ squares. For large numbers k and/or n this results in significantly different computation times; see for example Tables 1 and 2 in [44] for a direct comparison of the soft constrained approach (13) compared to the *ssq*-based AFS computation.

Finally, the AFS can be written as

$$
\mathcal{M} = \left\{ x \in \mathbb{R}^{s-1} : \ \min_S ssq(x, S) \le \varepsilon_{\text{tol}} \right\}
$$

for a fixed small parameter $\varepsilon_{\text{tol}} > 0$.

5.3. Geometric constructive feasibility test

The fundamentals of the geometric AFS construction are shortly outlined in Section 3.5.4. Principally, these

11

constructions are not limited to ($s = 3$)-component systems. However, the current literature does not contain any investigations for $s \geq 4$.

The geometric feasibility test of a certain point $x \in \mathcal{M}^+$ for the case $s = 3$ amounts to the following steps: First two tangents of INNPOL are constructed which run through the given point x and which (tightly) enclose INNPOL. Next the intersection of the first tangent with the boundary of $\mathcal{M}^+$ (the line segment between x and this point must touch at least one point w_i) is defined as P_1. The same is done for the second tangent. This results in the point P_2. Then x is a feasible point of the AFS if and only if the triangle with the vertices x, P_1 and P_2 includes the polygon INNPOL, see Section 3.5.1. An extension of this geometric construction which is applicable to noisy or perturbed data is developed in [22].

6. AFS computations for three-component systems

The definition of the AFS and the discussion of some of its numerous properties is now followed by a numerical algorithm for its computation. The focus is on the polygon inflation algorithm [43, 44] and its implementation in the *FACPACK* software. We also briefly discuss the geometric AFS construction, the triangle enclosure algorithm, the grid search approach and the MCR-Bands method for the computation of upper and lower band boundaries. We do not claim to present a complete discussion of all methods for AFS computations. For example, we do not consider the particle swarm algorithm for the detection of feasible regions [48].

6.1. Borgen plots and computational geometry

The geometric construction of Borgen plots has already been introduced in Section 3.5.4. This construction is purely geometric [6, 36]. The practical implementation of the construction on a computer requires methods of computational geometry. Hence a floating-point arithmetic is used so that certain approximation problems can occur. Perturbed spectral data, data which result from an SVD low rank approximation or spectral data containing small negative entries (e.g. from background subtraction) cannot successfully be treated with the classical Borgen plots. In [22] a generalized Borgen plot construction has been suggested which has extended the construction principles to perturbed spectral data. Generalized Borgen plots can be constructed with the *FACPACK* software.

6.2. Grid search

The grid search approach [50, 1] is a brute-force method to compute the AFS. It can be used for any $s \geq 2$. For two-component systems the function $f(x, S)$ is plotted on a proper grid in $\mathbb{R}^2$, see Section 4.1 or Figure 8. For the cases $s \geq 3$ one has to evaluate $\|f_0(x)\| + \min_S f(x, S)$ by Equation (13) or $\min_S ssq(x, S)$ on a suitable grid. Figures 11 and 10 show the function graphs for the data set 2 and for the data set 1. The grid search method can simply be implemented on a computer. However, the computational costs increase exponentially in the number of components s. Moreover, the grid search approach is a non-adaptive method. Any increase of the resolution by a factor κ (e.g., in each coordinate direction the number of grid points is multiplied by κ) results in an increase of the computational costs by the factor of κ^{s-1}.

6.3. Triangle enclosure

The triangle enclosure method was introduced in 2011 for three-component systems in [12]. The idea is to approximate the boundary of the two-dimensional AFS by series of equilateral triangles which cover the boundary. Therefore an initial triangle is computed for which at least one vertex x fulfills that $x \in \mathcal{M}$ and at least one vertex y is not located in $\mathcal{M}$. The third point z can be a feasible or a non-feasible point. Then the generation of further triangles starts. The idea is to mirror the most recently generated triangle along one of its edges which intersects the boundary of $\mathcal{M}$. This is done in a way that the triangle chain grows until the initial triangle is reached. Then the boundary of a segment of the AFS has been successfully approximated.

Figure 12 illustrates the idea of the triangle enclosure method. The shape of an Erlenmeyer flask is approximated by a chain of equilateral triangles. This is done for different edge lengths a. The precision of the boundary approximation is equal to the edge length a. Halving the edge length in order to increase the approximation quality doubles the number of required triangles.

For each segment (or isolated subset) of the AFS a separate chain of triangles is to be computed, see Section 3.5 for the possible cases. However, if the AFS consists of only a single topologically connected set, then this set contains a hole which encloses the origin. In this case an additional run of the triangle enclosure algorithm is required in order to approximate the inner boundary by a second chain of triangles. The computational costs of the triangle enclosure algorithm is significantly smaller compared to the grid search method. The reason is that feasibility tests are only required for points close to the boundary.
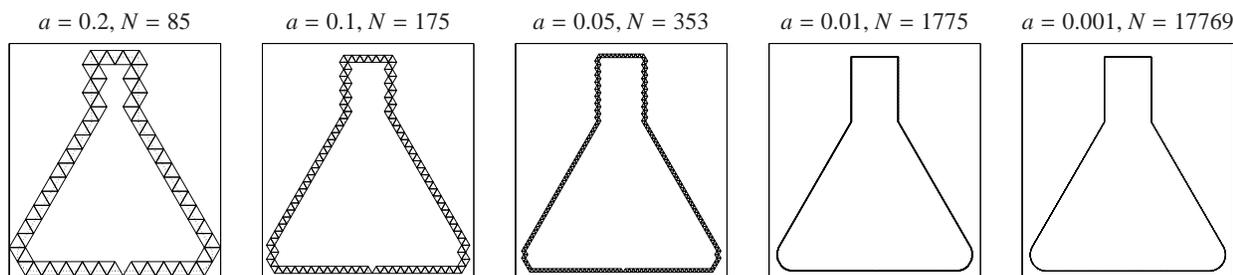
Figure 12: Approximation of the shape of an Erlenmeyer flask by the triangle enclosure method by chains of equilateral triangles with edge lengths a. The required number of triangles is N. The start triangle is chosen in the way, that its centroid is located in the middle of the bottom-line. The edge length a limits the precision of the boundary approximation.

6.4. MCR-Bands

The MCR-Bands method [10, 49, 21] is not generically an AFS computation method. Instead, MCR-Bands aims at the computation of lower and upper boundaries for the feasible bands of each component. The method works for any $s \geq 1$. The band boundaries are computed by minimizing (for the lower boundaries) or maximizing (for the upper boundaries) a certain cost function. As this method does not aim at a direct computation of the AFS we refer for further explanations to [49].

The minimal and maximal band boundaries show an interesting property. If a band boundary function for the possible spectra is expanded with respect to the first s right singular vectors, then the associated expansion coefficients (after FSV-scaling) are located on the boundary of the AFS; see e.g. [2] and [53] for systems with $s = 2$ or $s = 3$ components. So far, a systematic explanation for this has not been given and is a possible topic for future work. For noisy or perturbed spectral data, the localization of the band boundaries on the boundary of the AFS does not hold strictly. The reason is that the MCR-Bands toolbox and the feasibility check by Equations (12)–(15) deal in different ways with perturbed or noisy data.

6.5. Polygon inflation

The polygon inflation algorithm [43] and its algorithmic variation of inverse polygon inflation [44] are adaptive methods for the computation of the AFS for three-component systems. The idea is to approximate the boundary of each segment of the AFS by sequences of step-by-step refined polygons. A combination of a local error estimation with a local strategy for the refinement of the polygon results in a very effective, adaptive approximation scheme. Various control parameters allow to steer the approximation process and its quality. The polygon inflation method can be generalized to a

polyhedron inflation procedure in order to compute the AFS for four-component systems, see Section 7. We show some first examples in this work.

The geometric idea of the boundary approximation by inflating polygons is demonstrated in Figure 13. The shape of an Erlenmeyer flask is approximated in 2D by the polygon inflation method. Additionally the surface of a 3D Erlenmeyer flask is approximated by the polyhedron inflation algorithm. The surface of the polyhedron is a triangle mesh.

6.5.1. Steps of the polygon inflation algorithm

In this section the steps of polygon inflation method are explained. The inverse polygon inflation algorithm derives from the standard polygon inflation in a way that the sets $\mathcal{M}^+$ and $\mathcal{M}^*$ by Equations (7) and (8) are computed separately by inflating polygons. Then the AFS $\mathcal{M}$ is computed by forming the intersection $\mathcal{M}^+ \cap \mathcal{M}^*$. Thus the costs for the inverse polygon inflation is less than twice that of the polygon inflation. For further details see [43, 44]. Next the single steps of polygon inflation are explained for the case of an AFS consisting of three separate segments.

Step 1: Computation of an initial factorization of D

The first step of the polygon inflation method is to compute a nonnegative matrix factorization of D. According to Equations (2) and (5) this allows to find three points in the (spectral) AFS; the planar coordinates of these points are the matrix elements of the second and third column of T. If the AFS consists of three separated segments, then each segment contains exactly one of these points. For each of these three points an associated polygon is inflated.

The following steps 2 and 3 are executed for each of the three points from Step 1 which are located in the three AFS segments.
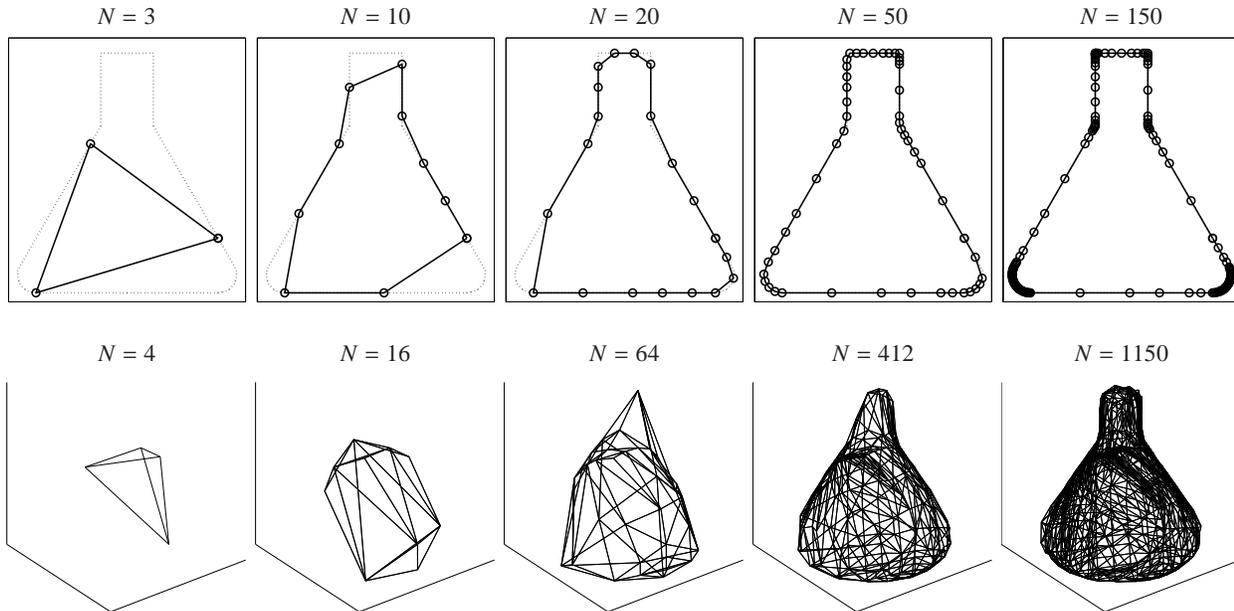
13

Figure 13: Top: Approximation of the shape of an Erlenmeyer flask in 2D by the polygon inflation algorithm with N vertices. Bottom: Approximation of a 3D Erlenmeyer flask by the polyhedron inflation algorithm. The surface of the polyhedron consists of N triangles.

Step 2: Computation of an initial polygon

The starting point is used in order to construct an initial triangle. The three vertices of this triangle are located on the boundary of the AFS segment. The resulting initial triangle includes the starting point.

Step 3: Step-by-step polygon inflation by edge subdivision

Given a current polygon with m vertices, a certain edge of this polygon is selected for a refinement. The selection rules are explained later in this section. If P_i and P_{i+1} are the endpoints of the selected edge, then the mid-perpendicular of the edge is determined. The closest point of intersection of this mid-perpendicular with the boundary of the AFS segment defines a new point P'_{i+1}. (There is always a second point of intersection of the mid-perpendicular with the boundary of the AFS. However, this more distant point on the opposite side of the AFS segment would not result in a successful approximation of the entire boundary by a polygon.) The vertices of the new and refined polygon read

$$[P_1, \ldots, P_i, P'_{i+1}, P_{i+1}, \ldots, P_m].$$

The introduction of the new vertex results in a gain or loss Δ_i of the area of the polygon. In most cases the area of the polygon increases, i.e. the polygon is inflated. If in rare cases the points P_i and P_{i+1} limit a non-convex region of the boundary of the AFS segments, then the refinement process can decrease the area of the polygon.

Figure 14 shows the three steps of the polygon refinement process for the data set 3. For the data set 2 with $w = 0$ the polygon inflation algorithm and the inverse polygon inflation algorithm are illustrated in Figure 15 by sequences of step-by-step refined polygons.

Polygon refinement

If a certain edge between the adjacent vertices P_i and P_{i+1} is marked for a refinement, then first the feasibility check, see Section 5.1, is applied to the mid-point $M = \frac{1}{2}(P_i + P_{i+1})$ of this edge. If M is feasible, then a second point $M' \notin \mathcal{M}$ is determined along the mid perpendicular of the current edge. The new point M' is to be found in the outward direction with regard to the current polygon. If M is not a feasible point, then a second point $M' \in \mathcal{M}$ is determined along the mid-perpendicular in the inward direction with regard to the current polygon. Hence the two points M and M' on the mid-perpendicular are determined. One of these points is a feasible point and the other point is not a feasible point. In other words, the boundary of the AFS segment intersects the mid-perpendicular between M and M'. The point of intersection is to be determined. This can be done in a slow but very stable way by means of the iterative bisection method. The iteration is stopped if a final precision ε_b has been reached.
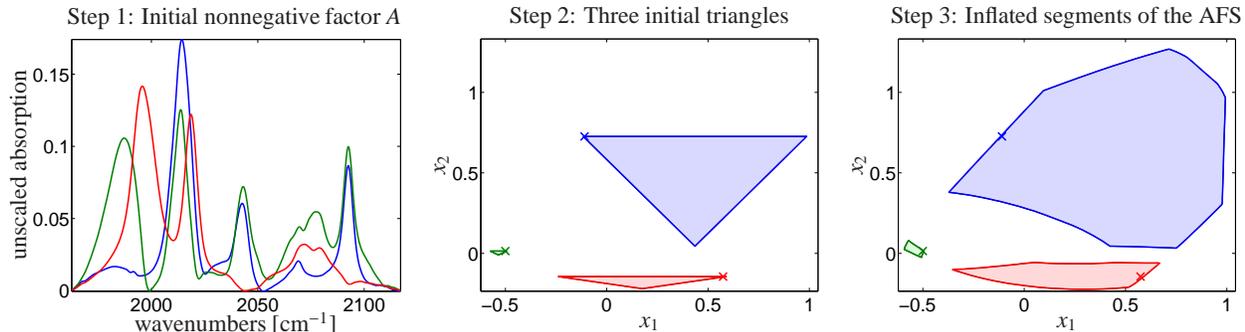
14

Figure 14: Data set 3: The steps of the polygon inflation procedure are illustrated. Left (Step 1): A nonnegative factorization of D results in a factor A (a so-called abstract factor). Center (Step 2): The initial factorization is represented by three points in the $x_1 - x_2$-plane of the AFS. These points are marked by $\times$ symbols. The three initial triangles for the inflation procedure are plotted. Right (Step 3): The inflation procedure (applied separately to the three initial triangles) results in a final AFS approximation. This reproduces the AFS shown in the second subplot of Figure 7.

Edge selection and stopping criterion for the polygon refinement

The change-of-area variable $\Delta_i \geq 0$ as introduced in Step 3 of the polygon inflation procedure can be stored as an attribute for each edge. If each of the three edges of the initial triangle has been subdivided at least once, then Δ_i has been assigned for each edge of the polygon. The stopping control parameter $\delta > 0$ is used to stop the polygon refinement process if $\Delta_j < \delta$ for all edges of the polygon. Then the gain or loss of area in the last subdivision of every edge is bounded by δ. If this stopping condition is not fulfilled, then an edge with the index ℓ is selected for a subdivision if $\Delta_\ell = \max_j \Delta_j$.

Figure 16 shows for the data set 2 with $w = 0$ a typical Δ_i convergence history in the *FACPACK* software. For these computations the stopping control parameter $\delta = 10^{-3}$ has been used.

6.5.2. The control parameters

The polygon inflation algorithm in the *FACPACK* software includes the following control parameters:

- $\varepsilon \geq 0$ is an upper bound for the acceptable relative size of negative entries in C and A, see Equations (12) and (13).
- ε_{tol} is the tolerance parameter in the feasibility test, see Equations (14) and (15).
- δ is the stopping parameter for the polygon refinement. The iteration is stopped if the nonnegative change of area attribute Δ_i of each edge of the polygon is bounded by δ. In the *FACPACK* software $\delta = 10^{-3}$ is the default value.
- ε_b is the control parameter on the guaranteed precision of the vertex localization close to the boundary of the AFS. This parameter should satisfy $\varepsilon_b \leq \delta$.

Suitable parameter selection for the parameter ε

The parameter ε on the acceptable relative size of negative entries in C and A considerably influences the size of the AFS. Thus ε should be selected carefully. A reasonable choice is as follows: If a computational procedure for nonnegative matrix factorizations (an NNMF routine) is only capable of producing a factorization $D = CA$ with matrices C and A with small negative entries, then ε can be selected according to

$$\varepsilon \geq \max\left(0, -\min_{i=1,\ldots,s} \frac{C(:,i)}{\|C(:,i)\|_\infty}, -\min_{i=1,\ldots,s} \frac{A(i,:)}{\|A(i,:)\|_\infty}\right). \quad (16)$$

As a rule of thumb, an increase of the value ε results in a growth of the area of the AFS. The crucial point is that increasing ε means that the factors C and A can contain (in absolute values) larger negative entries. Hence a larger number of approximate nonnegative matrix factorizations can be accepted as feasible. Thus the AFS grows. Figure 17 illustrates this relationship for the data set 3 for four different values of ε.

6.5.3. Benefit of the adaptive AFS approximation

A strength of the polygon inflation method is its adaptive strategy for the boundary approximation. If the boundary of the AFS is locally more or less a straight line, then only few vertices of the polygon are necessary for a high-quality boundary approximation. For other boundary regions which are strongly curved or non-smooth, a higher resolution of the polygon inflation method is a result of the local error estimation together with the local polygon refinement. Figures 12 and 13 demonstrate that only few vertices are sufficient for a good approximation of rectilinear regions of the boundary.

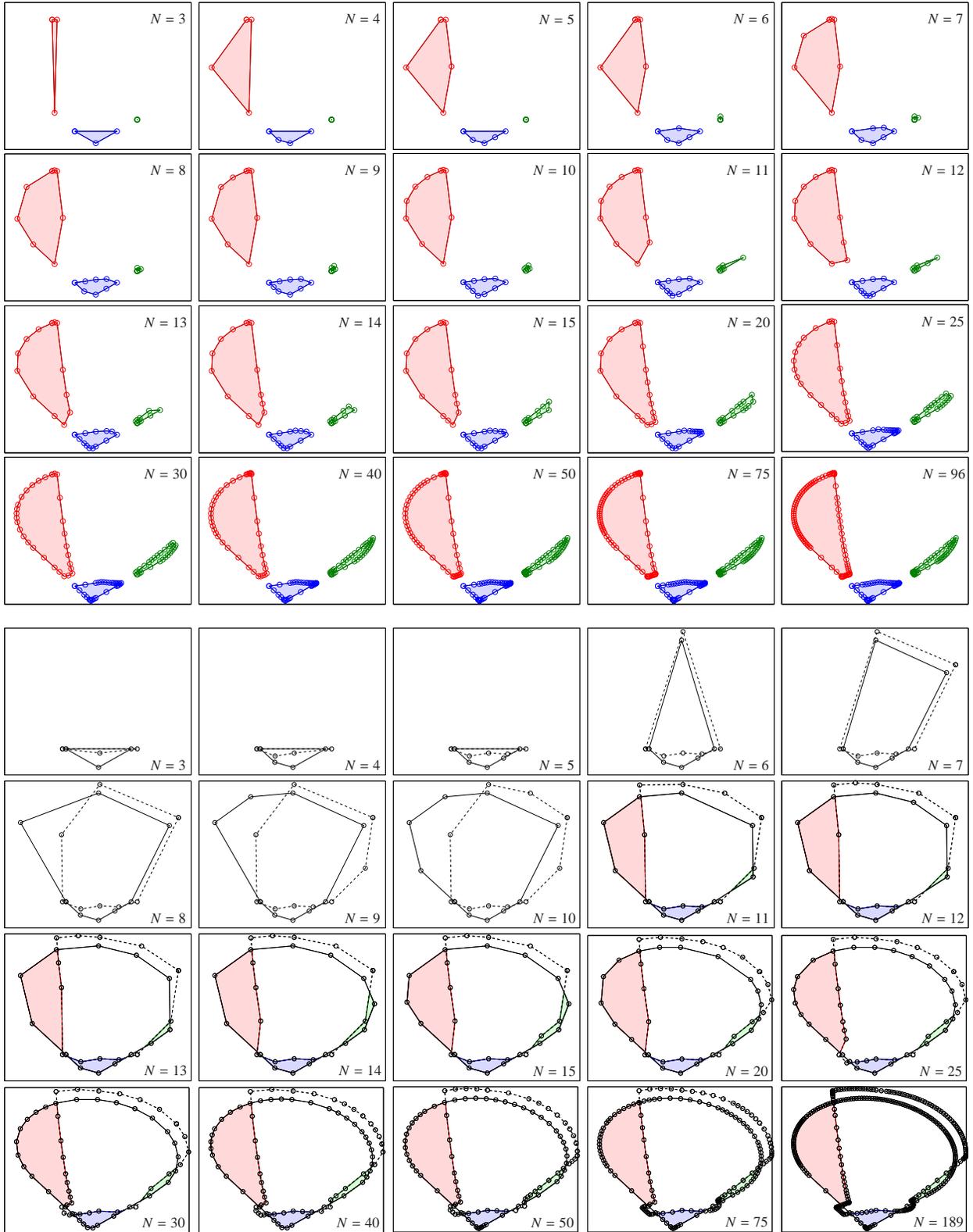Table 1 provides a comparison of the numbers of vertices which are necessary for an approximation with

Figure 15: Data set 2 with $w = 0$: Polygon inflation (rows 1-4) and inverse polygon inflation (rows 5-8) are illustrated for the same data matrix. In case of inverse polygon inflation the solid lines represent the boundary of $\mathcal{M}^+$ and the broken lines stand for the boundary of $\mathcal{M}^*$.
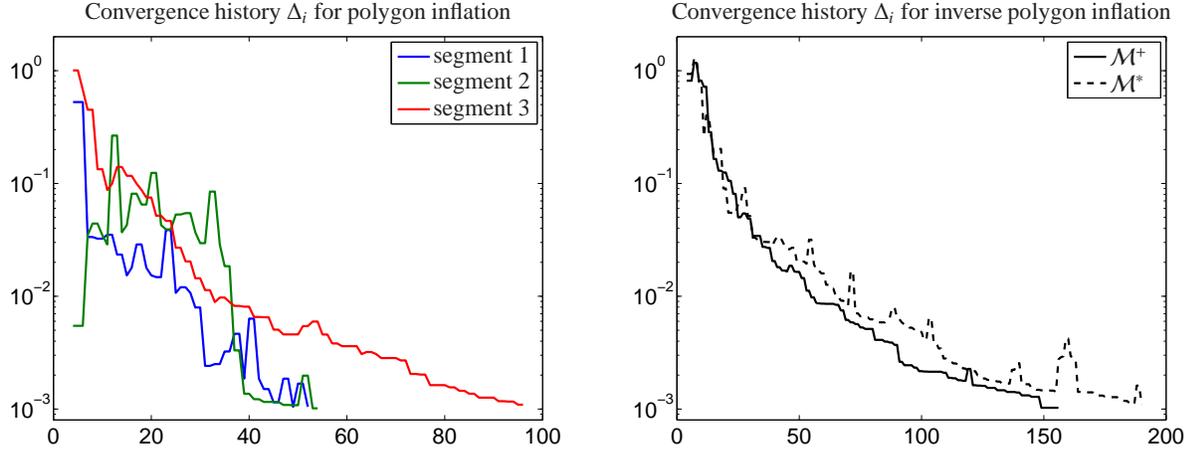
Figure 16: Data set 2: The gain/loss of area Δ_i of the adaptive polygon inflation procedure is plotted against the number of vertices of the polygon. Left: Δ_i for the three segments of the AFS. The polygon refinement is stopped if $\max \Delta_i \leq \delta$ with $\delta = 10^{-3}$. Right: Δ_i for the two polygons $\mathcal{M}^+$ and $\mathcal{M}^*$ for the case of the inverse polygon inflation procedure. (In the *FACPACK* software these convergence history data are contained in the log file "AFScomputation.log". Alternatively, one can push the log-file button.)

boundary precision	grid search	triangle enclosure	polygon inflation
10^{-1}	632	175	29
10^{-2}	$6.32 \cdot 10^4$	1775	85
10^{-3}	$6.32 \cdot 10^6$	17769	127
10^{-4}	$6.32 \cdot 10^8$	177735	299
10^{-5}	$6.32 \cdot 10^{10}$	1777349	901

Table 1: Approximation of the shape of an Erlenmeyer flask by the grid search approach, the triangle enclosure method and the polygon inflation method. Tabulated are the numbers of grid points (grid search approach), the total numbers of vertices of the approximating triangle chain and the numbers of vertices of the approximating polygons. The boundary precision values are $\varepsilon = 10^{-i}$, $i = 1, \ldots, 5$. These data affirm that the adaptive strategy has a clear advantage. See also Figure 12 and the first row in Figure 13.

guaranteed boundary precision of the shape of an Erlenmeyer flask by the grid search approach, by the triangle enclosure method and by the polygon inflation method. The boundary precision values are $\varepsilon = 10^{-i}$, $i = 1, \ldots, 5$. Further, Table 2 in [43] contains a similar comparison of the triangle enclosure method and the polygon inflation method for a three-component model problem.

7. AFS computations for four-component systems

All the methods for AFS computations for three-component systems as explained in Section 6 can be extended to four-component systems. The AFS for a four-component system is a bounded subset of the three-dimensional space. In this section a short overview is given on already published methods and on an on-going research project.

7.1. The slicing method

In [14] the idea has been presented to compute the three-dimensional AFS by two-dimensional slices. In this paper the triangle enclosure method has been used to approximate the boundary of the two-dimensional slices. For example, a slicing in z-direction works with a matrix T in (5) whose fourth column is fixed to a certain value ζ, i.e.

$$T(i, 4) = \zeta, \qquad i = 1, \ldots, 4.$$

(Additionally, the FSV scaling requires that $T(i, 1) = 1$ for all i.) All other matrix elements of T are free variables and are determined within the optimization procedure. With such a choice of T, one can compute the intersection of the 3D AFS with the 2D plane with $z = \zeta$. The entire AFS is computed by setting z to $m+1$ equidistant values in the interval $[z_b, z_u]$ with

$$\zeta_j = z_b + \frac{j}{m}(z_u - z_b), , \qquad j = 0, \ldots, m.$$

In each of the slices any of the AFS approximation methods from Section 6 can be applied.

Figure 18 illustrates the slicing method for the four-component data set 4. In each of the slices the polygon inflation method has been applied in order to compute the 2D intersections (the slices) with the 3D AFS. A
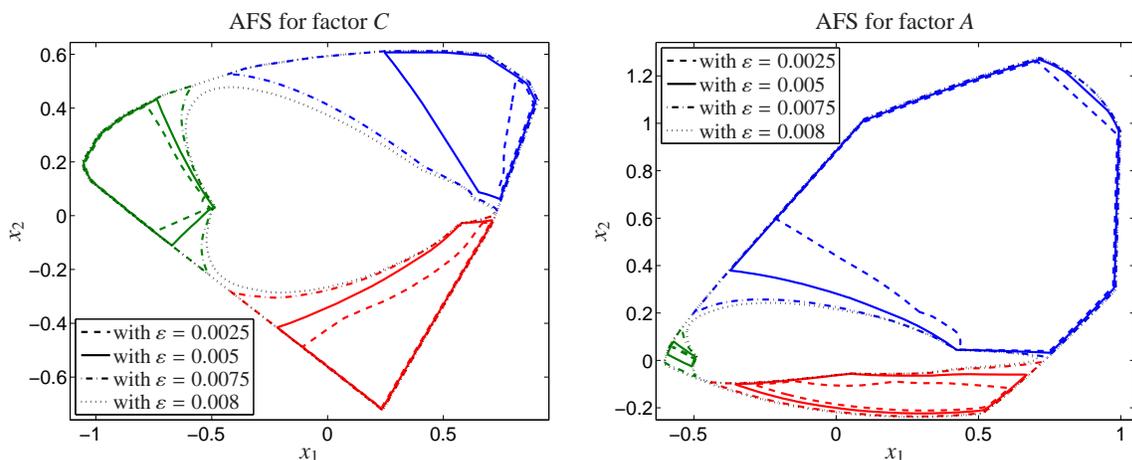
Figure 17: Data set 3: Variation of the control parameter ε on the acceptable relative size of negative entries in C and A, see Section 6.5.2. The area of the AFS segments increases with rising ε. For $\varepsilon \geq 0.008$ the initially isolated three segments are grown together into a single AFS with a hole (black dotted lines). The continua of feasible bands for the parameter $\varepsilon = 0.005$ are shown in the two rightmost plots of Figure 7.
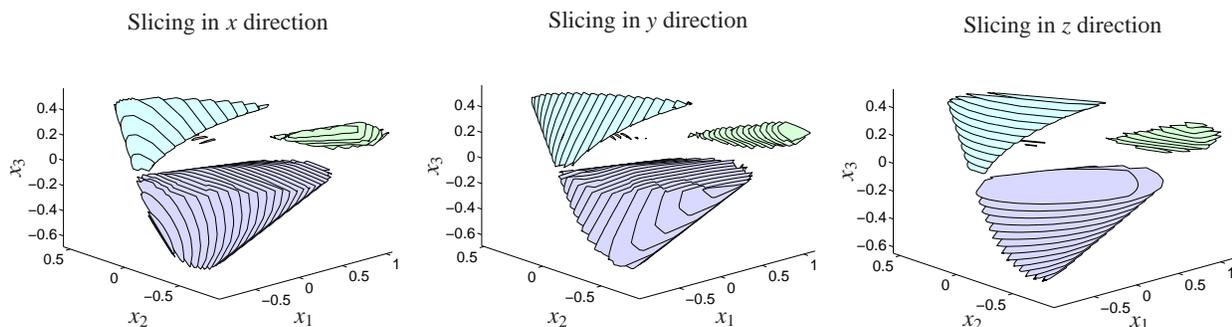

Figure 18: Data set 4 with $\sigma = 750$: Computation of the AFS for the spectral factor A by slicing along the three coordinate axes. A number of 30 slices is used for each plot. The same AFS computed by polyhedron inflation is shown in the right subplot of Figure 19.

number of 30 slices has been used for the slicing in the coordinate directions. The achievable resolution in the slicing direction depends on the number of slices.

7.2. Polyhedron inflation method

The direct generalization of the adaptive polygon inflation method is the polyhedron inflation method. To this end the 3D AFS is approximated by a series of polyhedrons. Each face of the polyhedron is a triangle. Thus the AFS is approximated by a 3D triangle mesh. The principles of the triangle subdivision and of the local error estimation are similar to those of the 2D AFS; but there are various technical challenges. Once again, adaptivity is an advantage of the inflation algorithm. We will explain the method in detail in a future work.

The method is demonstrated by the approximation of the shape of a 3D Erlenmeyer flask in Figure 13. Further, Figure 19 shows the concentrational AFS and the

spectral AFS for the four-component model system of the data set 4 with $\sigma = 750$.

8. Reduction of the rotational ambiguity by soft constraints represented in the AFS

The AFS construction includes the nonnegativity of the factors C and A as the one and only constraint. Sometimes additional information is available on the chemical reaction system. Such information is welcome in order to reduce the rotational ambiguity. For instance known concentration profiles or pure component spectra lead to a significant reduction of the rotational ambiguity by means of the duality- and complementarity theory [18, 34, 40, 4, 46, 17]. Alternatively, soft constraints can be very useful for extracting chemically meaningful solutions from the AFS [3, 47, 33].

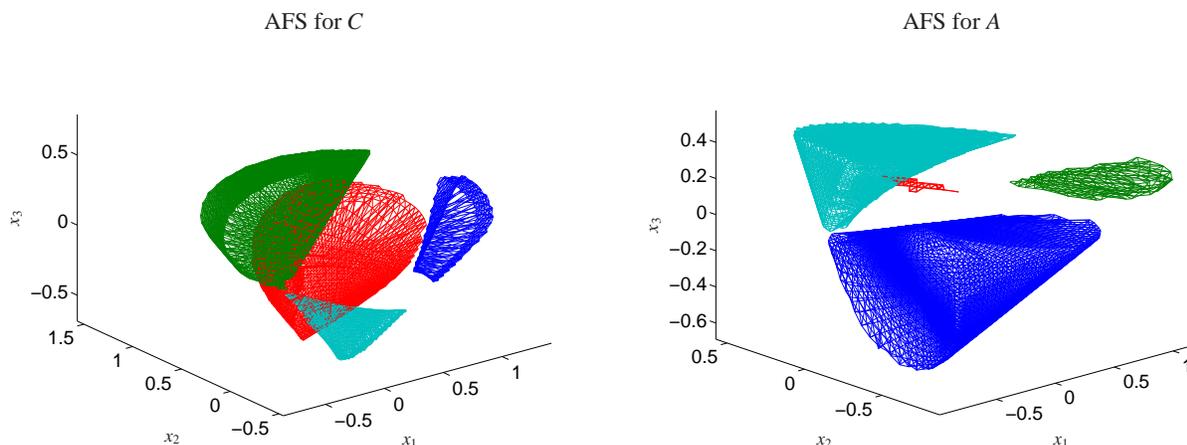AFS for C AFS for A

Figure 19: Data set 4 for $\sigma = 750$: For this four-component model problem the left plot shows the concentrational AFS and the right plot is the spectral AFS. The concentrational AFS is also shown in Figure 18.

8.1. Equality constraints: locked points in the AFS

The knowledge of a certain pure component spectrum is often called an equality constraint. It means that a certain point of the AFS is fixed (or locked). The effect on the remaining components is a reduction of the rotational ambiguity or equivalently a reduction of the area of the AFS segments. The reduced AFS can either be constructed geometrically (then one vertex in the simplex rotation algorithm [6] is fixed) or numerically (then a certain row of T is fixed). One observes that the reduction effect on the remaining segments of the AFS is relatively large if a point close to the origin is fixed.

All these concepts also apply to known concentration profiles. Then the whole procedure works with $D^T = A^T C^T$ where C and A have changed their places.

8.1.1. Complementarity theory and the AFS

The restrictions on the factor C which results from a given pure component spectrum are theoretically well understood by the complementarity theory or duality theory [34, 40, 4, 46, 31].

In general the AFS for an s-component system is a bounded subset of the $\mathbb{R}^{s-1}$. The complementarity theory provides a reduction of the rotational ambiguity to $(s-2)$-dimensional affine spaces. Theorem 3.1 in [46] proves that a given point x in the AFS for A forces the complementary concentration profiles represented by y to fulfill

$$\sum_{i=1}^{s-1} y_i x_i = -1.$$

Additionally locked points have an area-reduction impact on the AFS segment of the concentration fac-

tor of the same component for which the spectrum is given. All these restrictions can be combined if more than one point is locked. If s_0 points in the AFS for A are locked, then the admissible points y in the AFS for C for the complementary concentration profiles are in $s - 1 - s_0$-dimensional subspaces. This means for the important case of a three-component system that one known spectrum restricts the complementary two concentration profiles to a straight line in the AFS for C. Further the concentration profile of the chemical component whose spectrum is given is located in an AFS segment with a reduced area.

8.1.2. Equality constraints applied to experimental IR spectral data

Figure 20 illustrates for the data set 2 with $w = 0$ the reduction of the AFS for the cases of one or two given spectra. The left column shows the AFS sets for C and A only under the nonnegativity constraint. In the centered column the pure component spectrum of the component X is locked as a constraint (that is an equality constraint). The position of this spectrum is marked by a blue cross in the spectral AFS. This known spectrum results in a reduction of the areas of the AFS segments for the components Y and Z in the spectral AFS. Further the AFS segment for X in the concentrational AFS is reduced (black dashed lines). Due to the complementarity theory the AFS-segments for the components Y and Z in the concentrational AFS are reduced to linear AFS segments (black dashed lines). In the right column of Figure 20 a second spectrum is locked. The remaining segment in the AFS for A is reduced a second time. In the AFS for C a second line-restriction has been added.
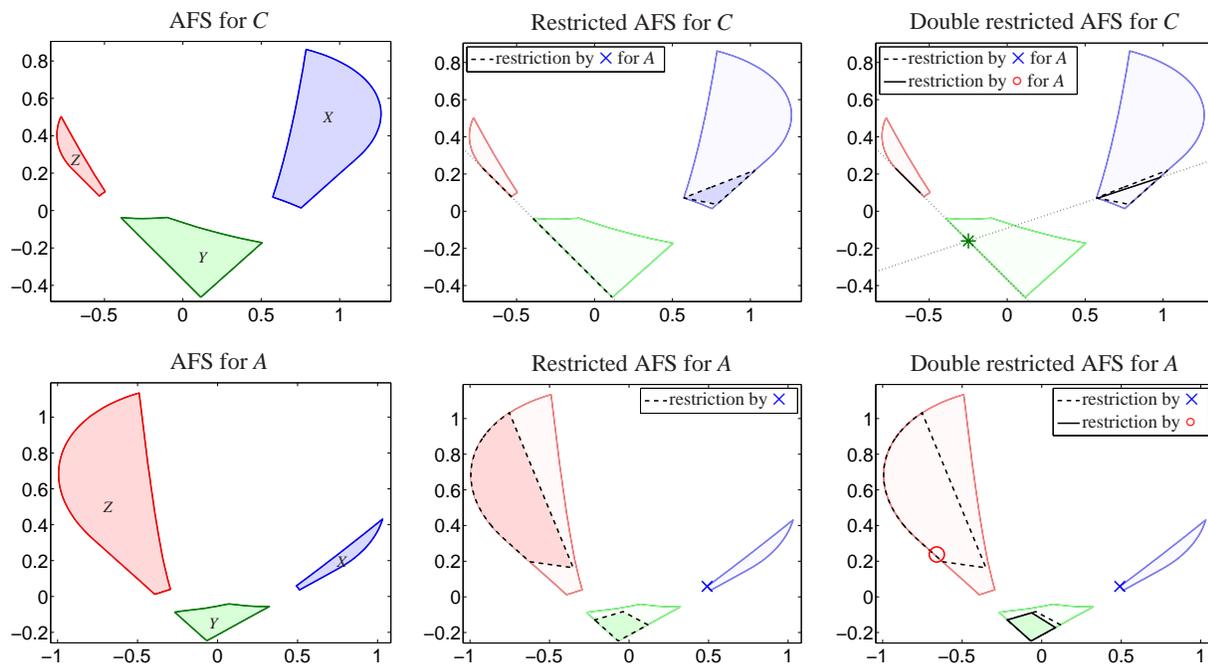
19

Figure 20: Data set 2: Application of equality constraints (this is the lock mode in *FACPACK*) to the AFS for *C* and the AFS for *A*. The three segments of the AFS sets and their associated chemical components *X*, *Y* and *Z* are shown in the left column.

Left column: Areas of feasible solutions without restrictions. Center column: The pure component spectrum of the component *X* is locked as a constraint. The position of the spectrum in the spectral AFS is marked by a blue cross. This known spectrum results in a reduction of the areas of the AFS segments for the components *Y* and *Z* in the spectral *A*. Further the AFS segment for *X* in the concentrational AFS is reduced (black dashed lines). Due to complementarity [40, 46] and duality theory [34], the AFS-segments for the components *Y* and *Z* in the AFS for *C* are reduced to linear AFS segments (black dashed lines). Right column: If additionally the pure component spectrum for component *Z* is locked (red ∘), then the segment for the component *Y* in the AFS for *A* is reduced a second time (solid line) and the concentration profile of *Y* is unique (green ∗ symbol). Finally, the reduced AFS segments for *X* and *Z* in the AFS for *C* are line segments (solid lines).

The point of intersection of these two lines uniquely determines the concentration profile of the component *Y* (green color).

8.2. Further soft constraints

The target function f by Equation (13) or the *ssq*-function in Section 5.2 can easily be extended by additional soft constraints in order to strengthen certain desired characteristics of the solution [4, 3, 33, 47]. In other words, one hopes to extract only those points from the AFS which additionally satisfy certain (soft) constraints. Finally and by using various constraints, one hopes to extract only one and chemically meaningful solution. Tauler [49] uses a similar approach within the MCR-Bands method.

Possible soft constraints are the unimodality constraint for the concentration profiles or even stronger a monotonicity constraint. Alternatively, one can use assumptions on closure or windowing, see Section 4.3 in [47]. Figure 21 for the data set 2 with $w = 0$ illustrates how the unimodality soft constraints can be used for a

reduction of the AFS. Additionally, Figure 21 shows the associated reductions of the feasible bands.

9. Geometric shapes of the AFS and AFS dynamics

The AFS can have various geometric shapes depending on the number s of independent components. For two-component systems the AFS always consists of two separated intervals. These intervals can degenerate to single-point intervals. One of these intervals contains only negative numbers and the other interval only positive numbers. For $s \geq 3$ the diversity of possible geometric shapes of AFS segments is much larger. In this section the focus is on the different (and partially degenerated) geometric shapes of the segments of an AFS. Furthermore the transition from an AFS with isolated segments to a one-segment AFS is demonstrated for model data sets with either three or four independent components.
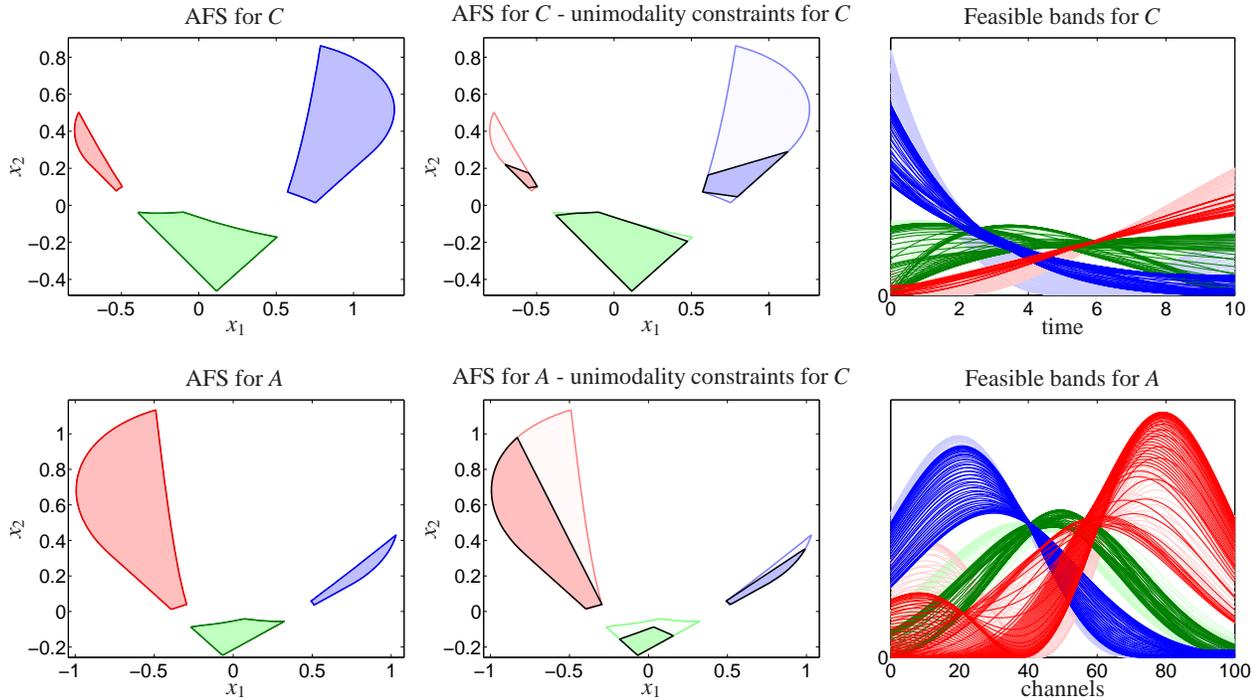
Figure 21: Data set 2 for $w = 0$: Application of the unimodality soft constraint.
Left column: AFS sets for C and A without restrictions. Center column: Application of monotonicity soft constraints only to the concentration profiles reduces the concentrational AFS and the spectral AFS. Right column: The associated feasible bands for C and A without unimodality soft constraint (pale colors) and with the unimodality soft constraint for the factor C (bright colors).

9.1. Shapes of AFS segments and degenerated segments

The segments of an AFS for an s-component system are subsets of the $\mathbb{R}^{s-1}$. Very often the $(s-1)$-dimensional volume of these segments is nonzero. However, certain problems can contain one or more unique spectra or concentration profiles. A further possible case is that the nonnegative matrix factorization problem for a given D has a unique solution [9, 25]. Such unique solutions are associated with single-point AFS segments. Further *degenerated cases* are line-shaped AFS segments in the case of 2D AFS sets for three-component systems. In the case of four-component systems the possible degenerated cases are single-point, line-shaped or planar AFS segments. These degenerated AFS segments have only been observed for model data. For perturbed or noisy experimental data the control parameter ε on tolerated negative entries inflates degenerated AFS segments, see Figure 17.

Computationally, single-point AFS segments can easily be computed. The initial nonnegative matrix factorization of D provides their coordinates. Somewhat more complicated is the computation of i_0-dimensional segments with $1 < i_0 < s - 1$. (E.g. for $s = 2$ the case

$i_0 = 1$ refers to a line-shaped AFS segment.) For these segments the polygon inflation algorithm (and also the triangle enclosure method), see Sections 6, as well as the polyhedron inflation method, see Section 7, are to be adapted properly. The start-up phase of the polygon inflation algorithm tries to build an initial triangle (with a nonzero area) and the polyhedron inflation algorithm tries to generate an initial non-degenerated tetrahedron, see left column of Figure 13. In other words, the algorithms aim at the construction of an s-simplex with a nonzero volume. If this is not successful, the algorithms try to construct an $(s-1)$-simplex with a nonzero $((s-1)$-dimensional) volume, and so on. The extreme and final case is that the algorithm ends in a 1-simplex, which is a unique point.

For three-component systems there is only the case of a line-shaped AFS segment between the extreme cases of a planar AFS segment with a nonzero area and a unique point. The numerical strategy for the computation of line-shaped AFS segments by polygon inflation is explained in Section 4.6 of [44]. The idea can easily be explained by means of a lighthouse. Starting from an initial feasible point x, which is considered as the position of the lighthouse, one determines a feasible beam
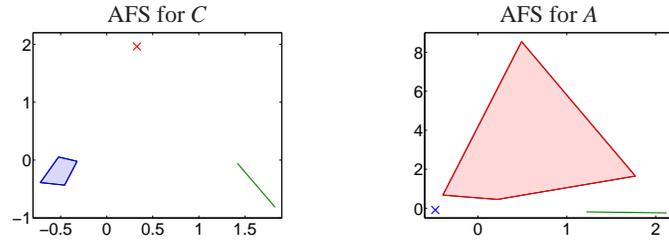
21

Figure 22: The AFS sets of the factors C and A for a three-component model problem. Each AFS consists of three isolated segments. One segment is a unique point (i.e. one spectrum and one concentration profile are uniquely determined), a further segment is line-shaped and the remaining third AFS segment has a geometric area larger than zero.
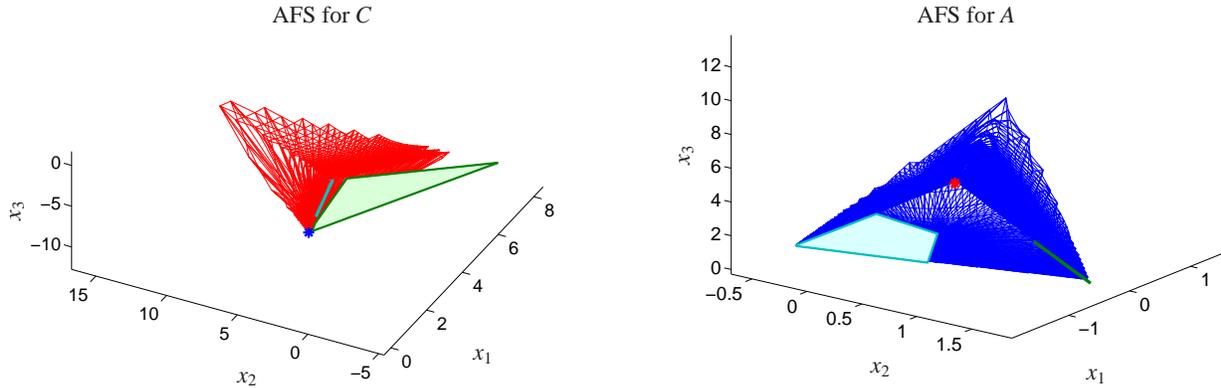

Figure 23: The AFS sets for a four-component model problem. One segment is a unique point (i.e. one spectrum and one concentration profile are uniquely determined), a second segment is line-shaped with a nonzero interval length, a third segment is planar with a nonzero planar area and the fourth segment is a volume segment with a nonzero volume.

direction so that the line

$$g_{r,x}(\varphi) = x + r \begin{pmatrix} \sin(\varphi) \\ \cos(\varphi) \end{pmatrix}$$

covers the desired line-shaped AFS segment. In a second step one determines the beam ranges in the form of bounds r_l and r_r so that $\mathcal{M} = \mathcal{L}_l \cup \mathcal{L}_r$ with

$$\mathcal{L}_l = \left\{ x + r \begin{pmatrix} \sin(\varphi - \pi) \\ \cos(\varphi - \pi) \end{pmatrix} \text{ with } r \in [0, r_l] \right\},$$

$$\mathcal{L}_r = \left\{ x + r \begin{pmatrix} \sin(\varphi) \\ \cos(\varphi) \end{pmatrix} \text{ with } r \in [0, r_r] \right\}.$$

This idea can easily be extended to the computation of planar AFS segments for the case of four-component systems; then two angles are to be determined.

Figure 22 shows for a three-component model system that the three shapes of AFS segments can occur simultaneously. Both the AFS for C and the AFS for A contain a single-point segment, a line-shaped segment and a planar segment with a nonzero area. Figure 23 illus-

trates the analogue of this for a four-component model problem.

9.2. AFS dynamics

Section 3.6 contains a discussion on the possible numbers of segments of an AFS. The obvious question on the inherent connection of this segment structure is discussed next. We demonstrate for the parameter-dependent data sets 2 and 4 that continuous changes of the parameter result in continuous changes of the AFS. Within this process of a dynamically changing AFS, the number of isolated segments of the AFS varies in a discontinuous way.

9.2.1. The three-component model data set 2

We compute for the data set 2 the AFS sets for a sequence of parameter values $w \in [100, 240]$. For $w = 100$ the signal width and thus the overlap of the pure component spectra is relatively small. The AFS for the spectral factor contains three separated segments, see Figure 24. For increasing w the segments show
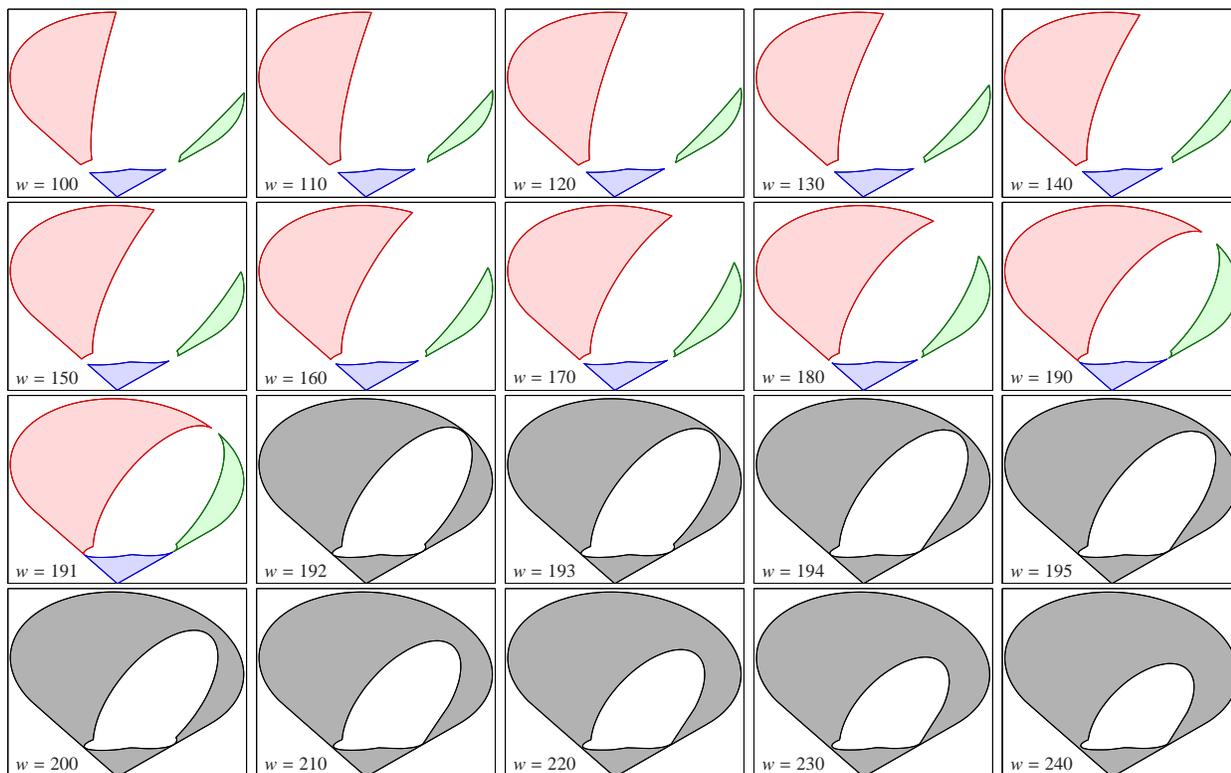
22

Figure 24: Data set 2: AFS dynamics with respect to a variation of the model parameter w. The shape of the spectral AFS is changing from three isolated segments to a one-segment AFS with a hole. An AFS with exactly two isolated segments is not possible.

an increasing geometric area; the distances between the segments decrease. For $w \geq 170$ the AFS computation by the *FACPACK* software should better be done by the inverse polygon inflation procedure. For $w \geq 192$ the AFS consists of only one segment, which contains a hole. The point of discontinuity w for which the three-segment AFS turns into a one-segment AFS is the same for the spectral AFS and the concentrational AFS.

9.2.2. The four-component model data set 4

Next the spectral AFS sets are computed for the model problem 4 for a sequence of parameter values σ with $\sigma \in [550, 850]$. The associated AFS sets are shown in Figure 25. For $\sigma = 550$ the AFS consists of four isolated segments. These segments grow together for increasing values of σ. Finally, for $\sigma = 810$ all four segments are connected and form a Swiss-cheese like 3D object with a complicated hole structure; the origin is contained in the hole.

Remark 9.1. *It is not clear by the results shown in Figure 25 if an AFS for a four-component system can consist of exactly two or exactly three segments. The res-*

olution of the computations is not high enough. This question has to be investigated analytically.

10. The FACPACK-toolbox for AFS-computations

The *FACPACK* toolbox [45] is a software package for AFS computations and can be applied to experimental spectroscopic data, to model data or even to any nonnegative matrix D. The software contains implementations of the polygon inflation method [44], the inverse polygon inflation method and the geometric constructive approach of generalized Borgen plots [22]. Moreover, it allows to apply the complementarity theory [46] within the AFS representation. The main functionalities are implemented in separate software modules. The current software revision is the *FACPACK* revision 1.2. Additionally, modules are contained for baseline corrections of spectroscopic data and for a library management of computed pure component spectra. The software can be downloaded on the *FACPACK*-homepage
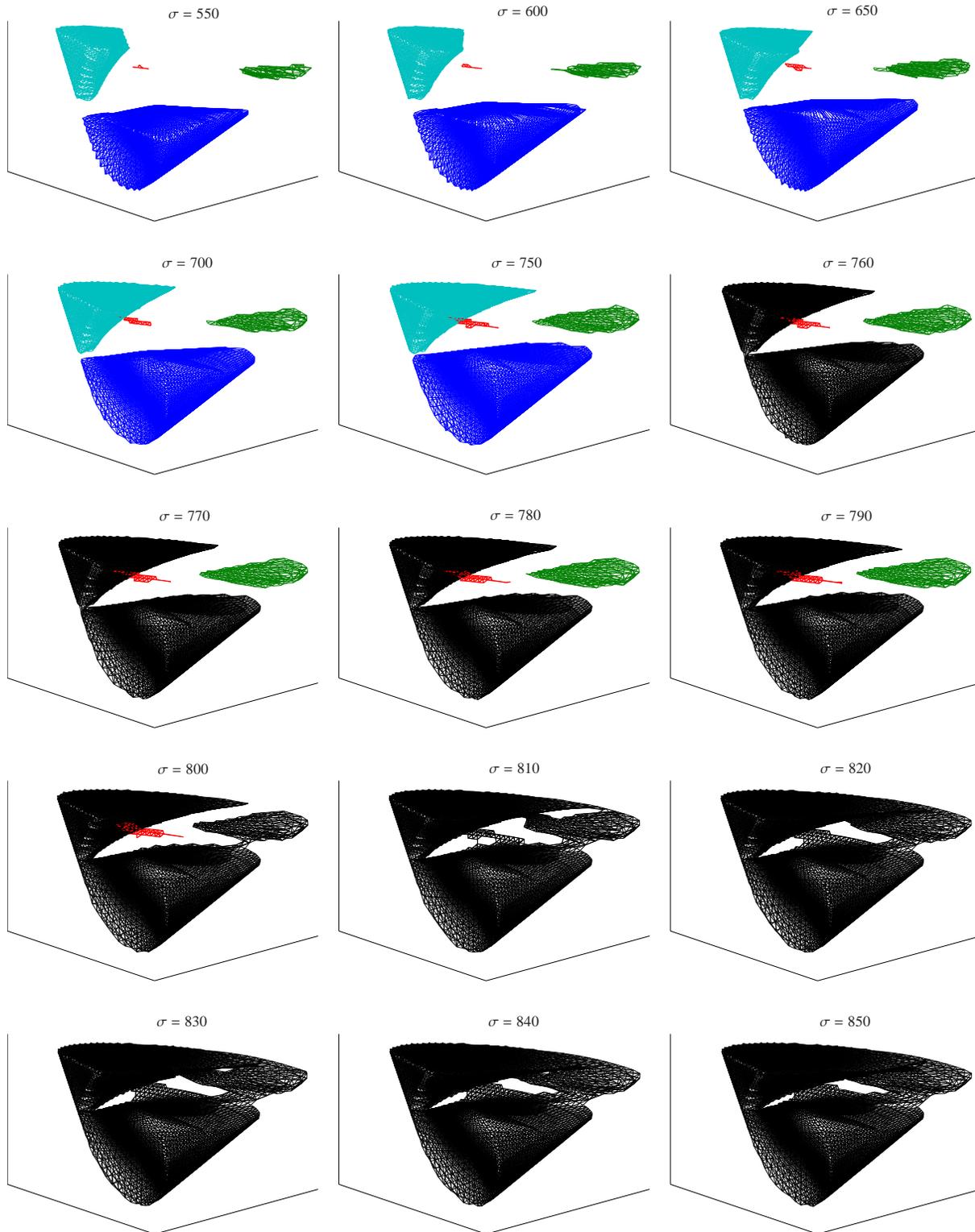
http://www.math.uni-rostock.de/facpack/.

23

Figure 25: Data set 4: AFS-dynamics with respect to a variation of the model parameter σ. For $\sigma = 550$ the spectral AFS contains four separated segments. Finally, for $\sigma = 850$ these segments are grown together to a one-segment AFS with a hole.

24

For more details on *FACPACK* see the manual [41]. This includes a guide on how to install and to start the software.

The *FACPACK* software comes with a graphical user interface in MATLAB. The time-consuming computational procedures are externalized in precompiled C-programs. Up to now the AFS-computation and the visualization is available for $s = 2$ and $s = 3$ component systems. For the future we plan to present an extension to four-component systems.

10.1. Data import

The spectral data matrix D is to be provided in a MATLAB compatible $\ast$.mat file and is assumed to contain row-wise the k spectra. Additionally, a vector of spectral wavenumbers/frequencies can be provided by an n-component vector variable x together with a vector t of time coordinate values.

10.2. The AFS computation module

The core of the *FACPACK* software is the polygon inflation algorithm for three-component systems. This software is contained in the *AFS computation*-module.

See Figure 26 for a screen shot of the *AFS computation*-module window with a short explanation of the possible users' operations. The first step in order to run the software is to push the button "AFS computation" in the start window. The second step is to load the data. The third step is to select whether a two- or three-component system is considered. The singular values $\sigma_1, \ldots, \sigma_4$ of D are displayed and serve indicators for the selection of the number of independent components. In step 4 an initial nonnegative matrix factorization is computed, which supplies the initial point for the construction of the initial triangles. Then, step 5 allows to select either the standard polygon inflation algorithm or the inverse polygon inflation algorithm. The AFS is computed by pressing the button "Compute AFS". A special feature of the software is its "live-view mode", see [45], which allows to move the mouse-pointer through the AFS and to see simultaneously the associated spectrum (or the concentration profile in case of the concentrational AFS).

The AFS computation module allows to apply equality constraints. Either one-, two- or three pure component spectra can be locked. Then the reduced AFS can be computed, see Section 8.1. External spectra (e.g. from a spectra library) can be imported; their coordinates are marked within the AFS.

10.3. The Complementarity & AFS module

Section 8.1.1 and in more detailed form the reference [46] explain the (possibly drastic) reduction of the rotational ambiguity by the complementarity theory. The module "Complementarity & AFS" illustrates this within the AFS.

The module is activated by pressing the button "Complementarity & AFS" in the start window of the *FACPACK* software. After data loading the third step is to compute the AFS sets for C and A. Alternatively, it is possible to compute only the set FIRPOL $\mathcal{M}^+$. The live-view mode allows to move the mouse-pointer through the AFS and to see simultaneously the associated spectra or concentration profile. By clicking the mouse button the current pointer position is locked and serves as a restriction for the complementarity theory. If three points either in AFS for C or A are locked, then a unique factorization $D = CA$ is defined. If some of these points are to be modified, then the user can repeat the steps 4-6. All steps are illustrated in Figure 27.

10.4. The Generalized Borgen Plot module

The *Generalized Borgen Plot*-module, see Section 6.1 for more details, allows the construction of the AFS for noisy of perturbed data by means of computational geometry methods.

First, the module is activated by the button "Generalized Borgen Plot" in the start window. Then the data is loaded and the desired scaling is selected. The further steps (AFS computation and live-view mode) are briefly explained in Figure 28.

11. Outlook and open problems

The AFS analysis which was launched by Lawton and Sylvestre [26] in 1971 and Borgen et al. [6, 5] in 1985 has produced a continuous stream of results for a deeper understanding of the rotational ambiguity of MCR methods. Especially, in the last decade a large number of papers have been published, see e.g. [34, 2, 12, 13, 43, 14, 44, 4, 3, 46, 22, 37].

Many challenges and open questions remain for the future. Some of them are:

1. How to integrate AFS methods to the toolboxes of practically working chemometricians? Up to now, AFS methods are more or less in the niche of research tools. AFS methods should be popularized as useful instruments for the investigation of the possible outcome of MCR methods.

2. How to compute and to visualize the AFS for systems with five or more independent chemical components? Then the AFS is a four- or even higher-dimensional bounded set whose graphical representation is very difficult. The derived representation by the dimension-independent feasible bands is a possible solution.

3. The nonnegative matrix factorization problem appears in various research fields as bio-informatics, text mining, neural sciences and others. It remains to be analyzed to which extent the global AFS approach could be helpful in order to investigate the inherent ambiguity for these factorization problems.

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 R. Tauler. Uniqueness and rotation ambiguities in Multivariate Curve Resolution methods. *Chemom. Intell. Lab. Syst.*, 108(2):100–111, 2011.

[3] S. Beyramysoltan, H. Abdollahi, and R. Rajkó. Newer developments on self-modeling curve resolution implementing equality and unimodality constraints. *Anal. Chim. Acta*, 827(0):1–14, 2014.

[4] S. Beyramysoltan, R. Rajkó, and H. Abdollahi. Investigation of the equality constraint effect on the reduction of the rotational ambiguity in three-component system using a novel grid search method. *Anal. Chim. Acta*, 791(0):25–35, 2013.

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

[6] 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.

[7] W. Chew, E. Widjaja, and M. Garland. Band-target entropy minimization (BTEM): An advanced method for recovering unknown pure component spectra. Application to the FT-IR spectra of unstable organometallic mixtures. *Organometallics*, 21(9):1982–1990, 2002.

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

[9] D. Donoho and V. Stodden. When does non-negative matrix factorization give a correct decomposition into parts? Advances in Neural Information processes, NIPS 2003, Cambridge, 2003.

[10] P.J. Gemperline. Computation of the range of feasible solutions in self-modeling curve resolution algorithms. *Anal. Chem.*, 71(23):5398–5404, 1999.

[11] P.J. Gemperline and E. Cash. Advantages of soft versus hard constraints in self-modeling curve resolution problems. Alternating least squares with penalty functions. *Anal. Chem.*, 75:4236–4243, 2003.

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

[13] A. Golshan, H. Abdollahi, and M. Maeder. The reduction of rotational ambiguity in soft-modeling by introducing hard models. *Anal. Chim. Acta*, 709(0):32–40, 2012.

[14] A. Golshan, M. Maeder, and H. Abdollahi. Determination and visualization of rotational ambiguity in four-component systems. *Anal. Chim. Acta*, 796(0):20–26, 2013.

[15] G.H. Golub and C.F. Van Loan. *Matrix Computations*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD, 2012.

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

[17] B. Hemmateenejad, Z. Shojaeifard, M. Shamsipur, K. Neymeyr, M. Sawall, and A. Mohajeri. Solute-induced perturbation of methanol-water association. *RSC Adv.*, 5:71102–71108, 2015.

[18] R.C. Henry. Duality in multivariate receptor models. *Chemom. Intell. Lab. Syst.*, 77(1-2):59–63, 2005.

[19] J. Jaumot, R. Gargallo, A. de Juan, and R. Tauler. A graphical user-friendly interface for MCR-ALS: a new tool for multivariate curve resolution in MATLAB. *Chemom. Intell. Lab. Syst.*, 76(1):101–110, 2005.

[20] J. Jaumot, P. J. Gemperline, and A. Stang. Non-negativity constraints for elimination of multiple solutions in fitting of multivariate kinetic models to spectroscopic data. *J. Chemometrics*, 19(2):97–106, 2005.

[21] 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.

[22] A. Jürß, M. Sawall, and K. Neymeyr. On generalized Borgen plots. I: From convex to affine combinations and applications to spectral data. *J. Chemometrics*, 29(7):420–433, 2015.

[23] C. Kubis, M. Sawall, A. Block, K. Neymeyr, R. Ludwig, A. Börner, and D. Selent. An operando FTIR spectroscopic and kinetic study of carbon monoxide pressure influence on rhodium-catalyzed olefin hydroformylation. *Chem.-Eur. J.*, 20(37):11921–11931, 2014.

[24] C. Kubis, D. Selent, M. Sawall, R. Ludwig, K. Neymeyr, W. Baumann, R. Franke, and A. Börner. Exploring between the extremes: Conversion dependent kinetics of phosphite-modified hydroformylation catalysis. *Chem. Eur. J.*, 18(28):8780–8794, 2012.

[25] H. Laurberg, M.G. Christensen, M.D. Plumbley, L.K. Hansen, and S.H. Jensen. Theorems on Positive Data: On the Uniqueness of NMF. *J. Chemometrics*, 2008, 2008.

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

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

[28] E. Malinowski. Factor analysis toolbox for Matlab. Applied Chemometrics, Inc., PO Box 100, Sharon, MA 02067, USA.

[29] C. Mason, M. Maeder, and A. Whitson. Resolving Factor Analysis. *Anal. Chem.*, 73(7):1587–1594, 2001.

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

[31] K. Neymeyr and Sawall. On an SVD-free approach to the complementarity and coupling theory: A note on the elimination of unknowns in sums of dyadic products. Technical Report, Accepted for J. Chemometrics, 2015.

[32] K. Neymeyr, M. Sawall, and D. Hess. Pure component spectral recovery and constrained matrix factorizations: Concepts and applications. *J. Chemom.*, 24:67–74, 2010.

[33] N. Rahimdoust, M. Sawall, K. Neymeyr, and H. Abdollahi. Investigating the effect of constraints on the accuracy of the area of feasible solutions in the presence of noise with an improved

cost function of the polygon inflation algorithm; soft versus hard constrained SMCR. Technical Report, Universität Rostock and IASBS Zanjan, 2015.

[34] R. Rajkó. Natural duality in minimal constrained self modeling curve resolution. *J. Chemom.*, 20(3-4):164–169, 2006.

[35] R. Rajkó. Additional knowledge for determining and interpreting feasible band boundaries in self-modeling/multivariate curve resolution of two-component systems. *Anal. Chim. Acta*, 661(2):129–132, 2010.

[36] 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.

[37] R. Rajkó, H. Abdollahi, S. Beyramysoltan, and N. Omidikia. Definition and detection of data-based uniqueness in evaluating bilinear (two-way) chemical measurements. *Anal. Chim. Acta*, 855:21 – 33, 2015.

[38] C. Ruckebusch and L. Blanchet. Multivariate curve resolution: A review of advanced and tailored applications and challenges. *Anal. Chim. Acta*, 765:28–36, 2013.

[39] M. Sawall, A. Börner, C. Kubis, D. Selent, R. Ludwig, and K. Neymeyr. Model-free multivariate curve resolution combined with model-based kinetics: Algorithm and applications. *J. Chemom.*, 26:538–548, 2012.

[40] M. Sawall, C. Fischer, D. Heller, and K. Neymeyr. Reduction of the rotational ambiguity of curve resolution techniques under partial knowledge of the factors. Complementarity and coupling theorems. *J. Chemom.*, 26:526–537, 2012.

[41] M. Sawall, A. Jürß, and K. Neymeyr. FAC-PACK: A software for the computation of multi-component factorizations and the area of feasible solutions, Revision 1.2. FAC-PACK homepage: http://www.math.uni-rostock.de/facpack/, 2014.

[42] M. Sawall, C. Kubis, E. Barsch, D. Selent, A. Börner, and K. Neymeyr. Peak group analysis for the extraction of pure component spectra. *Journal of the Iranian Chemical Society*, pages 1–15, 2015.

[43] 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.*, 27:106–116, 2013.

[44] 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.*, 28:633–644, 2014.

[45] M. Sawall and K. Neymeyr. *How to compute the Area of Feasible Solutions, A practical study and users' guide to FAC-PACK*, volume in Current Applications of Chemometrics, ed. by M. Khanmohammadi, chapter 6, pages 97–134. Nova Science Publishers, New York, 2014.

[46] 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.

[47] M. Sawall, N. Rahimdoust, C. Kubis, H. Schröder, D. Selent, D. Hess, H. Abdollahi, R. Franke, A. Börner, and K. Neymeyr. Soft constraints for reducing the intrinsic rotational ambiguity of the area of feasible solution. Technical Report, Accepted for Chemom. Intell. Lab. Syst., 2015.

[48] A.N. Skvortsov. Estimation of rotation ambiguity in multivariate curve resolution with charged particle swarm optimization (cPSO-MCR). *J. Chemometrics*, 28(10):727–739, 2014.

[49] 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.

[50] 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.

[51] E. Widjaja, C. Li, W. Chew, and M. Garland. Band target entropy minimization. A robust algorithm for pure component spectral recovery. Application to complex randomized mixtures of six components. *Anal. Chem.*, 75:4499–4507, 2003.

[52] W. Windig and J. Guilment. Interactive self-modeling mixture analysis. *Anal. Chem.*, 63(14):1425–1432, 1991.

[53] X. Zhang and R. Tauler. Measuring and comparing the resolution performance and the extent of rotation ambiguities of some bilinear modeling methods. *Chemom. Intell. Lab. Syst.*, 147:47–57, 2015.
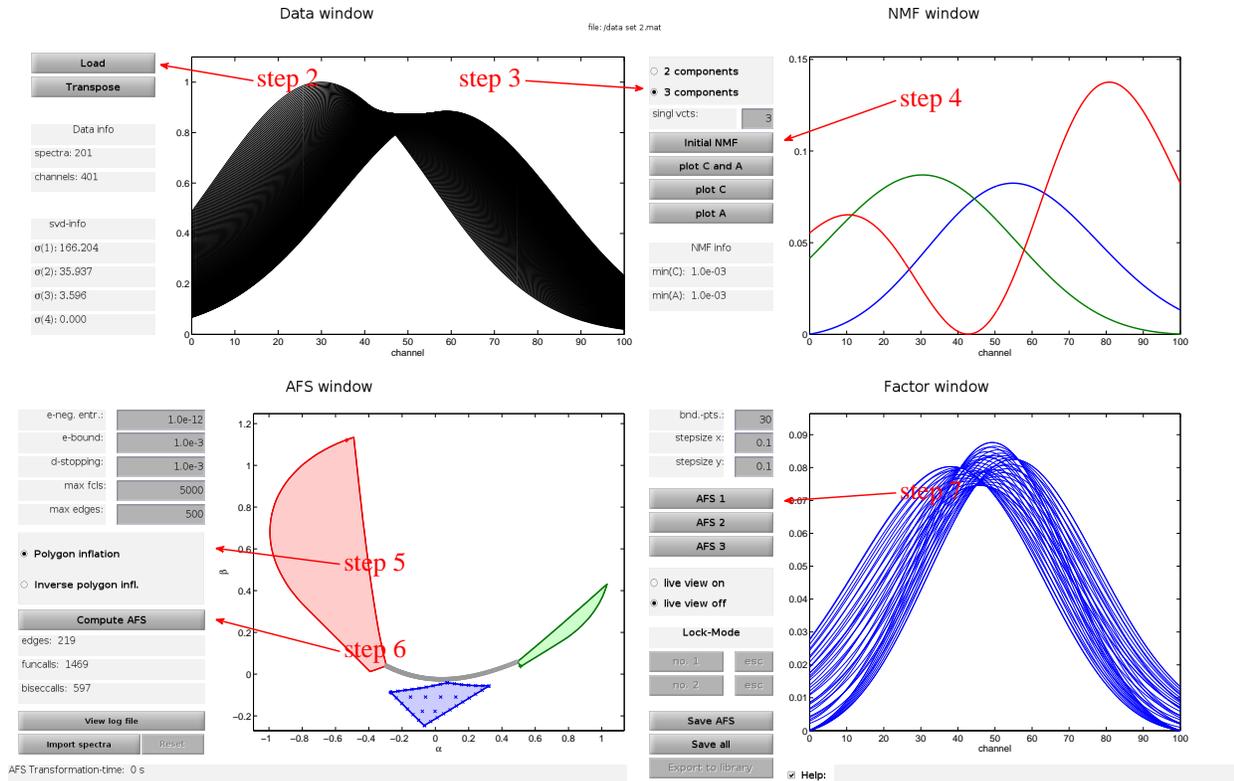
Figure 26: Quick-start in seven steps for the data set 2.

Step 1: Press the button **AFS computation** in the start window to activate this GUI.

Step 2: Load the data. Here data set 2 is used.

Step 3: Select 3 as the number of components.

Step 4: Compute an initial nonnegative matrix factorization (NNMF).

Step 5: Choose Polygon inflation.

Step 6: Compute the AFS which consists of three isolated segments.

Step 7: Plot the range of spectral factors (spectral bands), e.g., for the lowermost AFS segment.
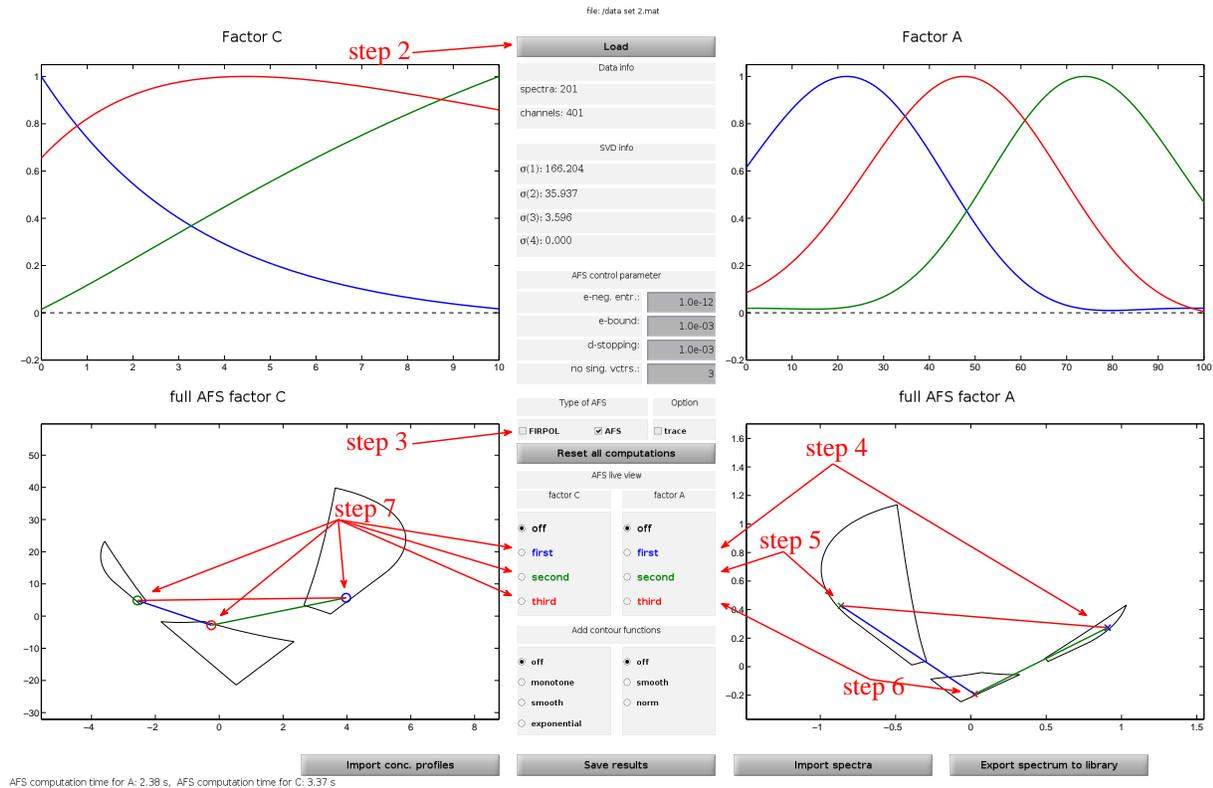
Figure 27: Complementarity theory and AFS implementation demonstrated for the data set 2.

Step 1: Press the button **Complementarity & AFS** to activate this GUI.

Step 2: Load the data. Here data set 2 is used.

Step 3: Either FIRPOL or the AFS sets for C and A can be computed. The computation of the AFS may be time-consuming.

Step 4: After clicking the button first, the mouse pointer can be moved through the spectral AFS. The associated spectra are plotted in the live-view mode. A certain spectrum $A(1, :)$ can be locked by clicking in the AFS.

Step 5: Click the button second to repeat the spectrum selection for a second spectrum $A(2, :)$. For the remaining third component the concentration profile is shown which is uniquely determined by the complementarity theory.

Step 6: A last third spectrum $A(3, :)$ can be selected by moving the mouse pointer through the spectral AFS. The resulting predictions on the spectral factor are shown interactively.

Step 7: These three buttons and also the buttons first, second, third can be clicked and then a spectrum or concentration profile can be modified by moving the mouse pointer through the AFS. This allows to modify the two triangles which represent a pure component factorization of the given spectral data matrix.
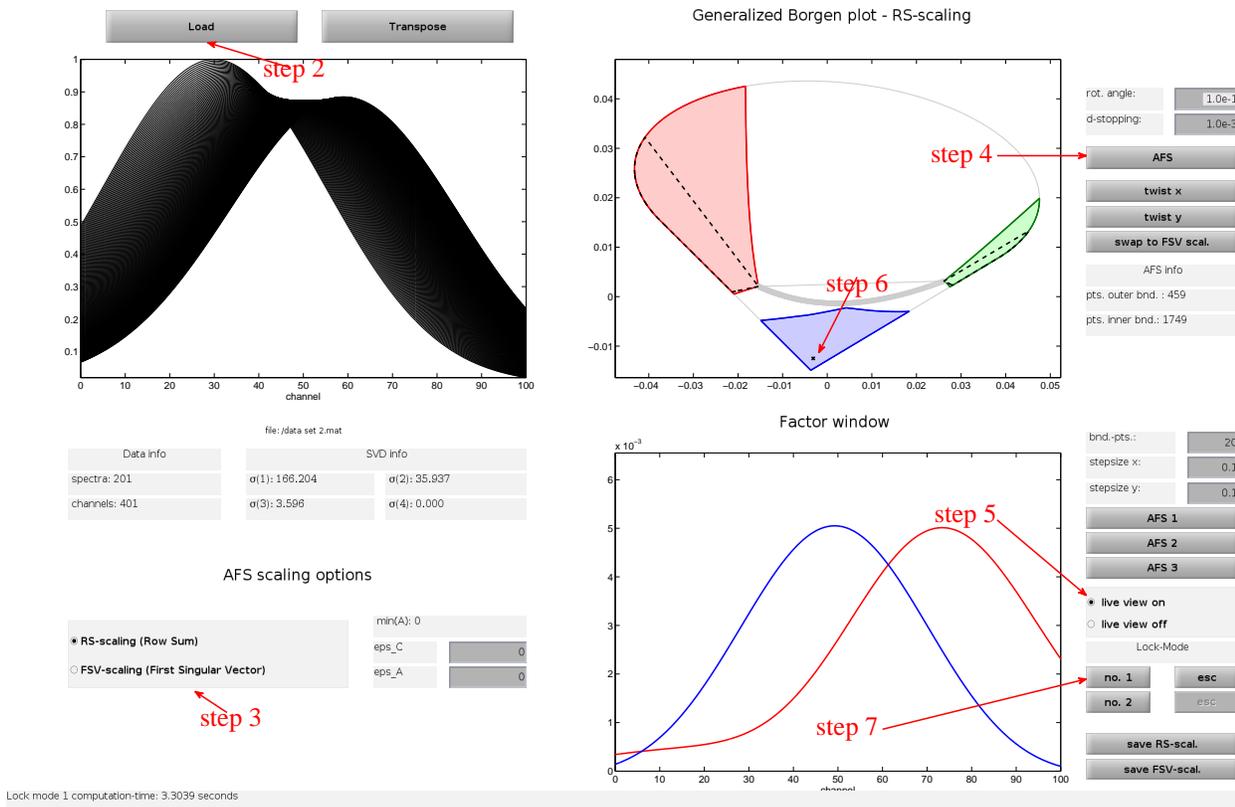
Figure 28: The Borgen plot module demonstrated for the data set 2.

Step 1: Press the button **Generalized Borgen Plot** to activate this GUI.

Step 2: Load the data. Here data set 2 is used.

Step 3: The scaling options are the Row Sum scaling and the First Singular Vector scaling.

Step 4: The AFS is computed and consists of three isolated segments.

Step 5: Activate the live-view on mode.

Step 6: Move the mouse pointer through the AFS and watch the interactively computed solutions. Points in the AFS can be fixed by clicking the left mouse button.

Step 7: If a point is fixed, then the reduced AFS segments for the remaining components are computed.