On the ambiguity underlying multivariate curve resolution methods.

Mathias Sawall[a], Henning Schröder[a], Denise Meinhardt[a], Klaus Neymeyr[a,b]

[a] *Universität Rostock, Institut für Mathematik, Ulmenstraße 69, 18057 Rostock, Germany.*
Email addresses: First name(dot)last name(at)uni-rostock.de, Telephone: +49 381 498 6640/6648.

[b] *Leibniz-Institut für Katalyse, Albert-Einstein-Straße 29a, 18059 Rostock, Germany.*

Abstract

Multivariate curve resolution (MCR) techniques can recover the pure component information from the spectral mixture data. Assuming an underlying bilinear model, the data matrix is to be factorized in matrices of the pure component spectra and the associated concentration profiles. Typically, there are solution continua exist which can be represented by the area of feasible solutions (AFS). This chapter analyzes the ambiguity of the matrix factorization problem. Various methods are reviewed for the construction and the numerical approximation of the nonnegative factorizations. An overview on the history of the AFS up to recent developments is included. Important properties of the AFS are discussed. Several example problems are analyzed mainly by the FACPACK software.

Key words: area of feasible solutions, FACPACK, multivariate curve resolution, nonnegative matrix factorization, spectral recovery.

1. Introduction

The spectroscopic analysis of the interaction of matter with electromagnetic radiation allows an extensive information retrieval not only from chemical reaction systems. Independently of the nature of the interaction of the radiation (mainly absorption, emission or scattering) the resulting spectra naturally result from superposition phenomena that are determined by the chemical components.

Typically, a series of spectra is recorded along the time axis of a chemical reaction system. In analogy to the *cocktail party problem* [1], or more generally the *blind source separation problem*, the aim is to extract the underlying pure component information from the spectral mixture data. In particular, one is interested in the number of the chemical species, their pure component spectra (in order to identify the chemical species) and their concentration profiles. Multivariate curve resolution (MCR) methods are prominent tools for unmixing the spectral mixture data. In favorable cases the data contains characteristic fingerprints that allow a direct identification of certain components. In most cases, MCR methods suffer from an intrinsic non-uniqueness of the pure component decomposition.

Here we consider bilinear data according to the Lambert-Beer law [2]. The spectral mixture data is stored in a $k \times n$ matrix D. The k measured mixture spectra are written row-wise in D, and the index n denotes the number of the spectral channels per spectrum. The matrix form of the Lambert-Beer law includes a time and a frequency dependence [3, 2, 4] and says that the matrix D (approximately) equals the product of the concentration profile matrix C and the spectra matrix S as

$$D = CS^T + E. \tag{1}$$

In its columns, the $k \times s$ matrix C contains the concentration profiles of the s pure components. Further, the columns of the $n \times s$ matrix S are the associated pure component spectra, and the error matrix E collects all (ideally small or vanishing) deviations from ideal bilinearity.

The spectral recovery problem means that the two matrix factors C and S are computed from the spectral mixture data D. Hence this problem is a *matrix factorization problem* with the important constraint that the factors C and S are nonnegative matrices. The nonnegativity constraint is not sufficient in order to guarantee that only one factorization exists aside from trivial permutations and rescalings of the columns of C and S. Instead, many nonnegative factorizations typically exist apart from the chemically correct factors. Mathematically, this fact is expressed in the existence of regular transformation matrices T so that

$$D = CS^T = C(T^{-1}T)S^T = (CT^{-1})(TS^T) \tag{2}$$

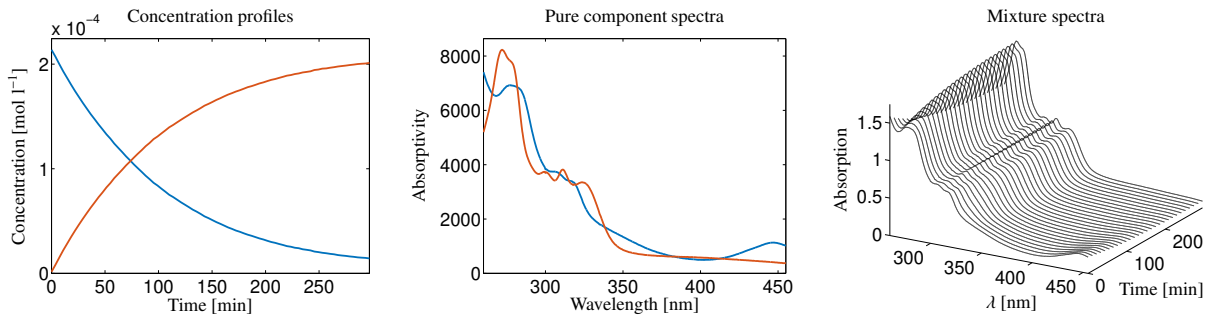

Figure 1: The concentration profiles (left), the pure component spectra (center) and the mixture spectra (right, only every third spectrum is plotted) for data set 1. The pure component spectra are scaled in a way that they are consistent with the concentration profiles by the kinetic reaction model and with $D = CS^T$. Blue (X): [Rh((R,R)-*i*Pr–ButiPhane)(COD)]BF_4, red (Y): [Rh((R,R)-*i*Pr–ButiPhane)($MeOH_2$)]BF_4.

with nonnegative matrices CT^{-1} and TS^T. These factors can represent a substantially different nonnegative factorization of D. This finding is well-known under the keyword of *rotational ambiguity* [5, 6, 7, 8]. A low-dimensional way to represent the sets of all possible columns of C and S that appear in nonnegative factorizations of D is the so-called *area of feasible solutions* (AFS). This chapter reports on the AFS, on some of its properties and on its numerical computation. The representation has a focus on the polygon inflation algorithm [9, 10] and its implementation in the FACPACK software package [11].

The chapter is organized as follows: First, in subsection 1.1 we introduce three data sets that accompany the explanations and serve to illustrate the various methods. In section 2 we give a brief review on MCR-approaches that use a truncated singular value decomposition (SVD) in order to construct C and S. The AFS is introduced in section 3. The presentation includes a short history of the AFS-approach and its geometric-constructive background with the relations to the so-called feasible simplices. Further, we discuss the duality principle and its interpretation in terms of the geometric AFS construction. Various methods for a numerical computation of the AFS are introduced in section 4. Section 5 is about the MatLab-toolbox FACPACK for AFS computations and related numerical algorithms. Finally, several results of numerical experiments are presented in section 6.

1.1. Data sets

In order to discuss the ambiguity underlying the spectral recovery problem, we use three spectroscopic data sets throughout this chapter. These spectral data sets are an experimental UV/Vis data set for a chemical two-component reaction system, a three-component model data set and an experimental three-component IR-data set. In this system a fourth major component exists, which does not show a significant absorption in the selected wavenumber interval. Each of the three data sets includes a kinetic reaction model, but it is not necessary to have a kinetic model for applying MCR methods. However, for representation purposes it appears to be favorable to use an accompanying kinetic model in order to construct the concentration profiles and so the model data. Furthermore, such a model helps to reduce the rotational ambiguity and to compare the reduction step with other approaches.

Data set 1. *This data set is on the investigation of ButiPhane ligands as well as the hydrogenation activity and selectivity of the corresponding rhodium complexes [12]. The data set covers a number of $k = 82$ UV/Vis spectra taken at equal time distances in the interval $t \in [0, 297]$min. Each spectrum uses $n = 1951$ equidistant spectral channels in the wavelength interval $\lambda \in [260, 455]$nm. The underlying kinetic reaction model is*

$$X \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} Y.$$

Therein X stands for [Rh((R,R)-iPr–ButiPhane)(COD)]BF_4 and Y for [Rh((R,R)-iPr–ButiPhane)($MeOH_2$)]BF_4. The initial concentration values are $c_X(0) = 2.15 \cdot 10^{-4}$ mol l^{-1} and $c_Y(0) = 0$ mol l^{-1}.

Figure 1 shows the concentration profiles of the two chemical components X and Y for the kinetic hard model solution $k_{-1} = 0\ \text{min}^{-1}$ and $k_1 = 9.494 \cdot 10^{-3}\ \text{min}^{-1}$, the associated pure component spectra as well as the mixture spectra forming the rows of D. The first four left and right singular vectors of D and a plot of the first 15 singular values are shown in figure 2.

Data set 2. *For the three-component model problem we use the kinetic reaction model*

$$X \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} Y \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} Z$$

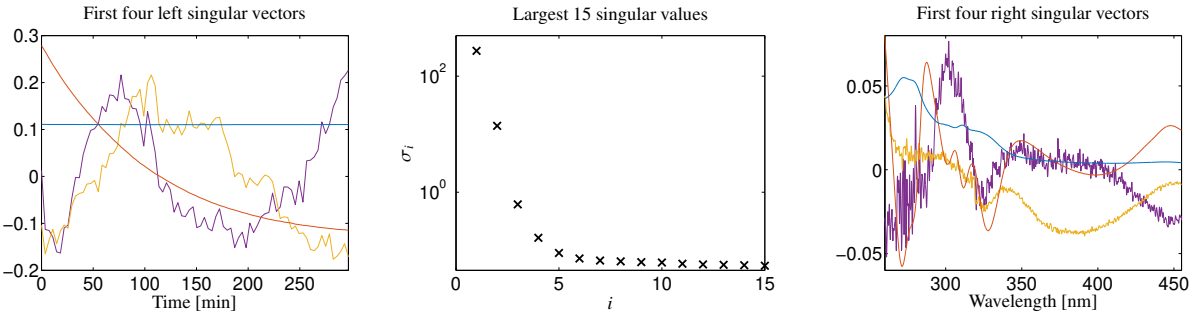

Figure 2: Results of the singular value decomposition (SVD, see section 2.3) of D for data set 1. Left: the first four left singular vectors. Center: the largest 15 singular values. Right: the first four right singular vectors. Colors in the order of an ascending count index: blue, red, ochre, purple.

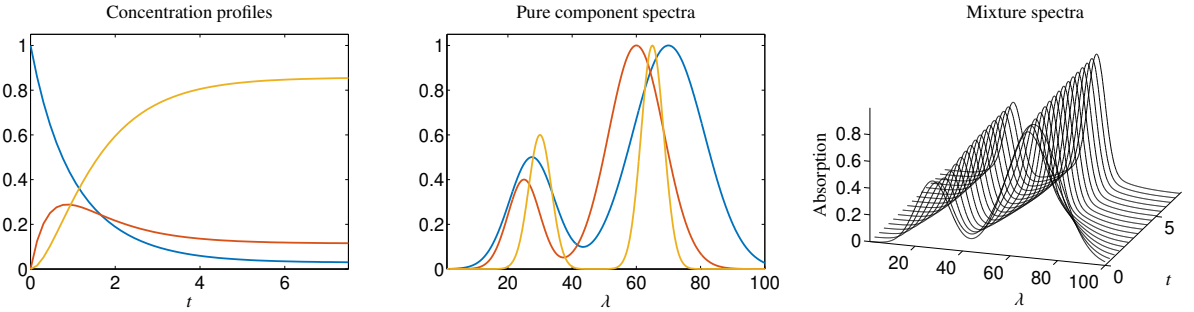

Figure 3: The concentration profiles (left), the pure component spectra (center) and the mixed spectra (right, only every third spectrum is plotted) for data set 2. Blue: X, red: Y, ochre: Z.

with $k_1 = 1$, $k_{-1} = 0.25$, $k_2 = 1.5$, $k_{-2} = 0.2$ *and the initial concentrations* $c_X(0) = 1$, $c_Y(0) = 0$ *and* $c_Z(0) = 0$. *The equidistant time interval discretization with* $t \in [0, 7.5]$ *uses* $k = 50$ *grid points. The computation of the concentration profiles of X, Y and Z that form the columns of C is done numerically with the* MatLab*-routine* ode45. *The* ode45 *control parameters are* RelTol= 10^{-10} *and* AbsTol= 10^{-10}. *The spectral profiles are assumed as sums of Gaussians in the following form*

$$a_X(\lambda) = 0.5 \exp\left(-\frac{(\lambda - 27.5)^2}{100}\right) + \exp\left(-\frac{(\lambda - 70)^2}{250}\right),$$
$$a_Y(\lambda) = 0.4 \exp\left(-\frac{(\lambda - 25)^2}{50}\right) + \exp\left(-\frac{(\lambda - 60)^2}{150}\right),$$
$$a_Z(\lambda) = 0.6 \exp\left(-\frac{(\lambda - 30)^2}{25}\right) + \exp\left(-\frac{(\lambda - 65)^2}{25}\right).$$

The wavelength interval $\lambda \in [1, 100]$ *is discretized equidistantly by using* $n = 500$ *nodes. The resulting columns of* S *are taken as* $S_{i1} = a_X(\lambda_i)$, $S_{i2} = a_Y(\lambda_i)$ *and* $S_{i3} = a_Z(\lambda_i)$ *for* $i = 1, \ldots, n$. *This results in the spectral mixture data matrix according to* $D = CS^T \in \mathbb{R}^{50\times500}$. *The pure component profiles as well as the spectral mixture data are presented in figure 3.*

Data set 3. *The data set was generated in order to analyze the influence of the carbon monoxide pressure on the rhodium-catalyzed olefin hydroformylation; see [13] for the details on this reaction. The data set comprises a number of* $k = 850$ *(non-equidistant in time) spectra, each with* $n = 645$ *equidistant frequency channels in* $\nu \in [1962.1, 2117.6]\,\mathrm{cm}^{-1}$. *The spectra are taken in the time interval* $t \in [4.72, 883.6]\,\mathrm{min}$. *The reaction starts at* $t = 0$. *A Michaelis-Menten kinetic*

$$S + K \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} [S - K] \xrightarrow{k_2} P + K \tag{3}$$

is assumed. The initial concentrations are $c_S(0) = 0.44364\,\mathrm{mol\,l}^{-1}$, $c_K(0) = 3.0440 \cdot 10^{-4}\,\mathrm{mol\,l}^{-1}$ *and* $c_{S-K}(0) = c_P(0) = 0\,\mathrm{mol\,l}^{-1}$. *The components S (the substrate/olefin component), K (the hydrido complex) and* $S - K$ *(the acyl complex) show significant absorptions in the given frequency window, but not P (the aldehyde).*

The pure component profiles (computed by kinetic hard-modeling [14, 15, 4, 16, 17]) as well as the mixture spectra are presented in figure 4. The dominant left and right singular vectors of the spectral data matrix are shown in figure 5.

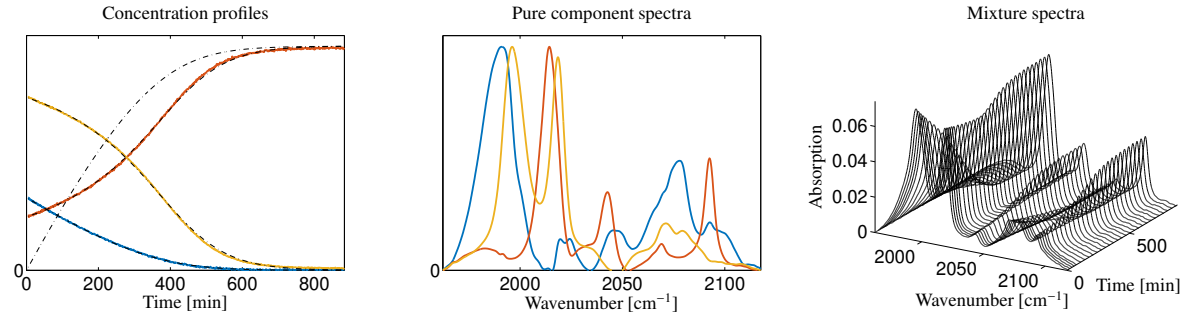

Figure 4: The concentration profiles (left), the spectral profiles (center) and the mixture spectra (right, only every 30th spectrum is plotted) for data set 3. Additionally, the solutions of the fitted kinetic model are drawn by dash-dotted black lines. Blue: olefin, red: hydrido complex, ochre: acyl complex.

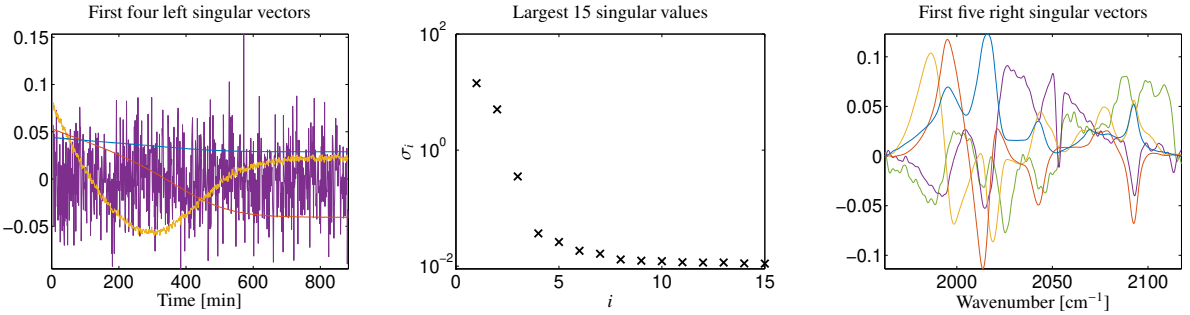

Figure 5: The dominant singular vectors and the 15 largest singular values of the spectral mixture data matrix D for data set 3. The fourth left singular vector is strongly oscillating so that no important chemical information can be expected from this singular vector. Colors in ascending order of the count index: blue, red, ochre, purple, green.

1.2. The colon notation

We use the so-called colon notation in order to get access to the columns and rows from a matrix, see [18]. If $M \in \mathbb{R}^{m\times n}$, then $M(:, i)$ is the ith column vector of M and $M(j, :)$ is the jth row vector of M. Furthermore, $M(:, i_1 : i_2)$ is the submatrix of M containing the columns i_1 up to i_2 of M. The colon notation is used by the numerical computing environment MATLAB.

2. Multivariate curve resolution

Multivariate curve resolution techniques try to reconstruct pure component information from spectral mixture data according to the bilinear Lambert-Beer law (1). Since nonnegativity of the factors C and S can be assumed, the mathematical problem underlying the MCR problem is that of a nonnegative matrix factorization (NMF, [19, 20]). A major complication for all MCR-methods and also for NMF-computation methods is the non-uniqueness of a factorization. In this section we analyze this non-uniqueness and discuss three strategies for the MCR-data analysis. The truncated SVD of D plays a key role not only for these methods but also for the construction of the AFS.

2.1. Ambiguity of the factorization

Three types of solution ambiguity are to be distinguished for the factorization problem $D = CS^T$. These are the permutation ambiguity, the intensity ambiguity and the rotational ambiguity [7].

The *permutation ambiguity* expresses the fact that the columns of C and S in a nonnegative factorization $D = CS^T$ can simultaneously be rearranged by inserting a permutation matrix P and its inverse matrix $P^{-1} = P^T$. Permutation matrices are formed by permuting the columns (or rows) of an identity matrix. Then

$$D = CS^T = C(PP^T)S^T = (CP)(P^TS^T) = (CP)(SP)^T \tag{4}$$

yields the factors CP and SP. The matrix CP has the same column vectors as C, but in a different order. The same is true for SP and S. The simultaneous rearrangement of the columns does not provide any deeper insight on the pure

components, but justifies to focus only on the possible first columns of C and S if the set of all feasible factorizations of D is considered. The *intensity ambiguity*, also called the multiplicative ambiguity [4], is similar to the permutation ambiguity. The only difference is that the permutation matrix P is substituted by a diagonal matrix with positive diagonal elements and that P^T is substituted by the inverse of the diagonal matrix. In other words, if a column of C is scaled by constant $\gamma > 0$, then the associated column of S is scaled by $1/\gamma$. Clearly, the scaling ambiguity does not provide any additional information. Hence we ignore these two ambiguities and also their combinations.

The only non-trivial ambiguity is the so-called *rotational ambiguity*. Then (4) changes to (2) with a non-diagonal, regular matrix T. If CT^{-1} and TS^T are nonnegative matrices, which is by no means always true, then a significantly different pure component decompositions of D has been determined. Typically, many such matrices T exist. Hence, not only a unique factorization does exist, but a continuum of possible nonnegative factorizations. One says that the MCR problem suffers from the rotational ambiguity. The associated spectral recovery problem has the aim to compute not only one of the many possible factorizations, but the chemically true pure component decomposition.

2.2. How to deal with the rotational ambiguity

MCR methods have different strategies how to deal with the (disturbing effect) of the rotational ambiguity. Some of these strategies are as follows:

1. Apply soft constraints in the form of regularization terms. This so-called soft modeling uses assumptions on the smoothness of the solution, on its unimodality or monotonicity, on the area-integral of the spectrum and so on. The intention is to steer the factorization process in a direction that finally yields the chemically correct solution.
2. Use a certain hard modeling. The solution must strictly satisfy this constraint. An important hard constraint is the consistency of the concentration profiles of the pure components with a given kinetic model. Typically, optimal reaction rate parameters are determined in the computational process.
3. Compute the full range of all possible columns of the factors C and S that allow a reconstruction of the given spectral data matrix D in the form $D = CS^T$. One set represents the possible concentration profiles of the pure components and the other set represents the corresponding pure component spectra. These sets of profiles can either be drawn in the form of so-called feasible bands [21] or in terms of its expansion coefficients with respect to the bases of singular vectors. The latter form is the AFS representation.

Several computational methods are available for each of these three approaches:

First, if soft constraints are used, then the pure component factorization can be computed by a numerical minimization of an objective function that depends on the factors C and S. Then the objective function minimizes the reconstruction error, namely that $D - CS^T$ should be close to the zero-matrix. The objective function further contains a weighted mean of various penalty and regularization functions. These additional terms of the objective function work in the sense of a Tychonoff regularization. Alternative ways to apply constraints in MCR methods are iterative schemes like nonnegative least-squares optimizations which can be combined with update formula in order to implement the constraints. Typical soft constraints are the smoothness, the unimodality, the monotonicity, sharp peaks and small entropy. Popular methods are, among others, the famous MCR-alternating least squares method (MCR-ALS, [22, 23]), the band target entropy minimization (BTEM, [24, 25]), the peak group analysis (PGA, [26]) or even classical NMF algorithms ([27, 28]), which have no chemistry-related background.

Second, if a chemical reaction is investigated for which a kinetic reaction model is available, hard-modeling is the method of choice [29, 14, 15, 4, 17]. Then an objective function is constructed that numerically computes the concentration profiles by the solution of an initial value problem for the given ordinary differential equations. The reaction rate constants are not assumed to be known. They are fitted in a way that CS^T optimally approximates D in the least squares sense. Popular methods that can include hard models are MCR-ALS, SPECFIT [29] or ReactLab [30].

Third, we consider the global approach to compute, instead of a single solution, the set of all possible profiles or their low-dimensional representatives. This process does not include any soft- or hard-modeling. The approach to compute extremal profiles by the MCR-BANDS software [31] is closely related, see also [32, 21]. MCR-BANDS explicitly computes two profiles for each component. For $i = 1, \ldots, s$ one either minimizes the *signal contribution function*

$$g_i(R) = \frac{\|C(R^{-1})(:,i)R(i,:)S^T\|_F^2}{\|D\|_F^2} \tag{5}$$

or maximizes the function with respect to regular matrices $R \in \mathbb{R}^{s\times s}$. The nonnegativity constraint is always active. The extremal profiles $C(R^{-1})(:,i)$ and $R(i,:)S^T$ provide rough estimates on minimal and maximal profiles. More

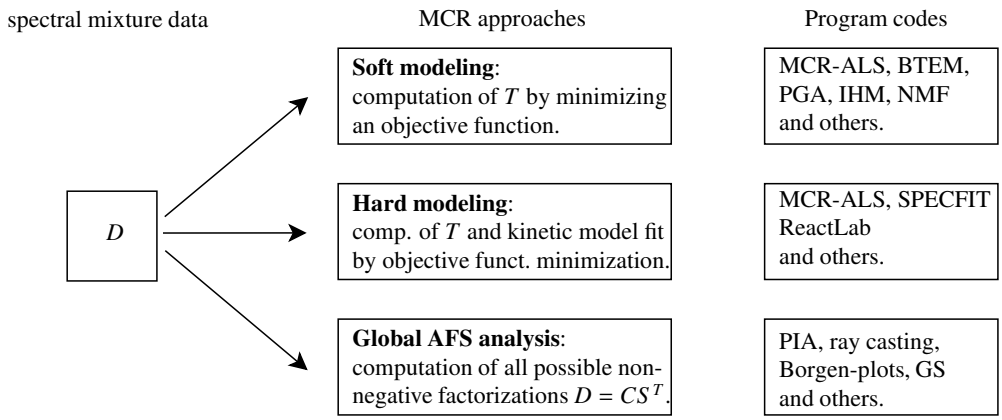

Figure 6: Spectral mixture data analysis by MCR-methods. The acronyms are as follows (only one main reference is given for each method): MCR-ALS: multivariate curve resolution by alternating least squares [23]. NMF: nonnegative matrix factorization [28]. BTEM: band target entropy minimization [25]. PGA: peak group analysis [26]. IHM: indirect hard modeling [33]. PIA: polygon inflation algorithm [9]. GS: grid-search [34]. References for SPECFIT, ReactLab and ray casting are [29], [30] resp. [35].

fundamentally, the AFS approach produces the full range of all feasible solutions and is explained in a detailed way in the following section.

An overview on the listed MCR-approaches and some popular methods is presented in figure 6.

2.3. The singular value decomposition

The singular value decomposition (SVD), see [18], of the spectral mixture data matrix D is the key-ingredient for the reconstruction of the factors C and S. The SVD of D reads

$$D = \tilde{U}\tilde{\Sigma}\tilde{V}^T$$

with orthogonal matrices $\tilde{U} \in \mathbb{R}^{k\times k}$ and $\tilde{V} \in \mathbb{R}^{n\times n}$. Further $\tilde{\Sigma} \in \mathbb{R}^{k\times n}$ is an all-zero matrix except for its diagonal elements. The columns of $\tilde{U}$ are the left singular vectors and the columns of $\tilde{V}$ the right singular vectors. The diagonal elements of $\tilde{\Sigma}$ are the singular values $\tilde{\Sigma}_{ii} = \sigma_i \geq 0$ with a decreasing order $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{\min(k,n)}$. For a matrix D of the rank s it holds $\sigma_s > 0 = \sigma_{s+1} = \ldots = \sigma_{\min(k,n)}$. In general, the number of the nonzero singular values indicates the number of independent absorbing species. The *independence* of species refers to the linear independence of the vectors which represent the concentration profiles of the pure components and the independence of the associated pure component spectra. However, for noisy and perturbed data one typically observes several (positive) singular values that are close to zero. In ideal cases there is a significant difference between those large singular values that represent chemical species and the other small singular values that are induced by noise.

The *truncated SVD* ignores the small singular values and the associated singular vectors that do not carry significant chemical information. Let s be the number of relevant singular values and let

$$U = \tilde{U}(:, 1:s), \qquad \Sigma = \tilde{\Sigma}(1:s, 1:s), \qquad V = \tilde{V}(:, 1:s)$$

be the matrices of the dominant singular vectors and singular values. Then

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

is a truncated SVD of D by matrices U, Σ and V which each have the rank s. The Eckhart-Young theorem [36, 37] quantifies the lack-of-fit of the truncated SVD, which is a low rank approximation of D, as $\|D - U\Sigma V^T\|_2 = \sigma_{s+1}$ in terms of the spectral norm $\|\cdot\|_2$. See also [38, 39] on these low-rank approximations in the context of the principal component analysis.

The truncated SVD provides the bases of the dominant left- and right singular vectors in order to construct the pure component factors in the following form

$$C = U\Sigma T^{-1}, \qquad S^T = TV^T, \tag{6}$$

see for example [5, 3, 4, 40]. The effect of (6) is a reduction of the variables of the pure component recovery problem. The point is that the two matrices C and S have a total of $s(k + n)$ matrix elements. However, it is ineffective to use

these matrix elements as the variables for minimizing the reconstruction error $D - CS^T$ under the constraint $C, S \geq 0$. Instead, equation (6) allows to form C and S only by using the s^2 matrix elements of T.

It is worth mentioning that MCR methods can also work in an SVD-free form. The alternating iterative methods for nonnegative matrix factorizations [27, 28] and the alternating least-squares curve resolution method MCR-ALS [22, 23] belong for instance to the SVD-free methods. However, the latter methods can still use an SVD for pre-filtering the data. The point is that an SVD of the spectral data matrix can be used in order to form a low-rank approximation with its rank being equal to the number of expected chemical components. Then the SVD-free factorization method can be applied to the filtered data.

3. The AFS as a global solution approach

As derived in the last section, the s^2 matrix elements of T are sufficient in order to represent feasible, nonnegative matrices C and S. Next we reduce these s^2 degrees of freedom to the smaller number of only $s-1$ degrees of freedom. The $s-1$ degrees are sufficient in order to represent all possible nonnegative columns of either C or S in terms of the AFS.

The justification for this reduction is that the set of all possible columns of the factor C (mutatis mutandis of the factor S) in all nonnegative pairs of factors $C = U\Sigma T^{-1}$ and $S = VT^T$ is equal to the set of all possible *first* columns in these factorizations. This enormous simplification derives from the permutation ambiguity, namely that any column order is possible; see equation (4). In other words, we are interested in the following two sets of the feasible concentration profiles and feasible spectra that can appear in a pure component decomposition of D

$$\mathcal{C} = \left\{ C(:,1) \in \mathbb{R}^k : \text{ exists invertible } T \in \mathbb{R}^{s\times s} \text{ with } C = U\Sigma T^{-1} \geq 0,\, S = VT^T \geq 0 \right\},$$
$$\mathcal{S} = \left\{ S(:,1) \in \mathbb{R}^n : \text{ exists invertible } T \in \mathbb{R}^{s\times s} \text{ with } C = U\Sigma T^{-1} \geq 0,\, S = VT^T \geq 0 \right\}.$$

The decisive first columns of C and S are completely determined by the first row of T, respectively the first column of T^{-1}. Hence, we have traced back the representations of these two sets to the s matrix elements in the first row of T (first column of T^{-1}). Next, we use the intensity ambiguity that allows us to assume a certain scaling for the single rows of T respectively the columns of T^{-1}. We select a scaling which makes the first component of the first row vector of T equal to 1. Alternatively, we can do this for the first component of the first column of T^{-1}. A mathematical justification for these scalings is required (since leading zero components would prevent a scaling). Fortunately, the Perron-Frobenius spectral theory of nonnegative matrices [41, 42] guarantees that the first singular vectors necessarily contribute to any nonnegative linear combination of the singular vectors. This provides the justification to assume that the first column of T is equal to $(1, \ldots, 1)^T$; for more details see section 3.4 or Theorem 2.2 in [10]. Hence, the low-dimensional representation of the set $\mathcal{S}$ works with the matrix T in the form

$$T = \begin{pmatrix} 1 & x_1 & \cdots & x_{s-1} \\ 1 & & & \\ \vdots & & W & \\ 1 & & & \end{pmatrix}. \tag{7}$$

Therein, W is an $(s-1)$-by-$(s-1)$ matrix which is to be specified later. Let $x = (x_1, \ldots, x_s)^T$ be the column vector of the matrix elements in the first row of T. On the basis of these settings we define the AFS for the spectral factor as

$$\mathcal{M}_S = \left\{ x \in \mathbb{R}^{s-1} : \text{ exists } W \in \mathbb{R}^{(s-1)\times(s-1)} \text{ in (7) such that rank}(T) = s,\; C = U\Sigma T^{-1} \geq 0 \text{ and } S = VT^T \geq 0 \right\}. \tag{8}$$

In words, $\mathcal{M}_S$ contains all $x \in \mathbb{R}^{s-1}$ so that a regular matrix T exists with $T(1,:) = (1, x^T)$, which results in nonnegative factors $C = U\Sigma T^{-1}$ and $S^T = TV^T$. The analogous definition of the AFS for the concentration factor reads (in a somewhat shortened form)

$$\mathcal{M}_C = \left\{ y \in \mathbb{R}^{s-1} : \text{ exists a regular } T \in \mathbb{R}^{s\times s} \text{ with } (T^{-1})(:,1) = (1, y^T)^T \text{ and } C, S \geq 0 \right\}. \tag{9}$$

3.1. The underlying reduction of the degrees of freedom

The *conditio sine quo non* for representing the inherent ambiguity of the MCR problem is a reduction of the degrees of freedom of the factors. This reduction process is summarized in this section. First of all it is worth noting that the AFS-sets do not depend on the (large) dimensions k and n of the data matrix D. The first reduction step comes from

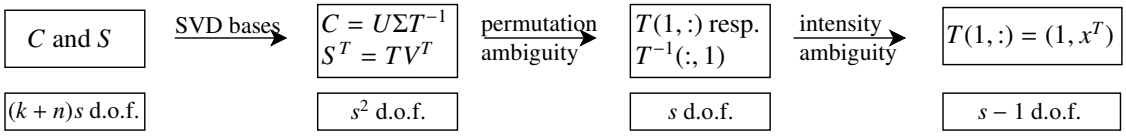

Figure 7: Reduction of the degrees of freedom (d.o.f.) underlying the low-dimensional representation of feasible profiles of the AFS.

the column-basis representation introduced in (2) or equivalently on the SVD-based representation (6). In this step the $s(k+n)$ matrix elements of C and S are represented in terms of the s^2 matrix elements of the matrix T. Typically, it holds $s \ll k, n$ for spectroscopic data.

The focus on the first columns of C and S, as justified above, reduces the representation of the rotational ambiguity to s variables namely to the first row of T respectively the first column of T^{-1}. To explain this important reduction step of s^2 degrees of freedom of T to only s degrees of freedom, we consider the following nonnegative factorization of D

$$D = \underbrace{U\Sigma T^{-1}}_{=C}\,\underbrace{TV^T}_{=S^T}.$$

Further let $\mathcal{P}$ be the permutation matrix formed by interchanging the first and the ℓth column of the $s \times s$ identity matrix. Then

$$D = \underbrace{U\Sigma T^{-1}}_{=C}\,\underbrace{TV^T}_{=S^T} = U\Sigma(T^{-1}\mathcal{P})\,(\mathcal{P}^T T)V^T = (C\mathcal{P})\,\mathcal{P}^T S^T = (C\mathcal{P})\,(S\mathcal{P})^T$$

shows that the permutation $S\mathcal{P}$ of the columns 1 and ℓ of S corresponds with the permutation $\mathcal{P}^T T$ of the rows 1 and ℓ of T. Similarly, the permutation $C\mathcal{P}$ of the columns 1 and ℓ of C corresponds with the permutation $T^{-1}\mathcal{P}$ of the columns 1 and ℓ of T^{-1}. In other words, if a certain pure component profile appears in a nonnegative factorization of D, the same profile appears in another nonnegative factorization in the first column of C or S after a permutation. Finally, the scaling assumptions $T_{11} = 1$ for the spectral AFS and $(T^{-1})_{11} = 1$ for the concentration factor AFS reduces the number of variables to $s - 1$. Thus the AFS-sets are subsets of the $\mathbb{R}^{s-1}$ with the rank s of D. These reduction steps are illustrated in figure 7.

3.2. The inner and outer polygons and their relations to the AFS

The geometric interpretation of feasible solutions and also the geometric construction of the AFS-sets rest upon two important polygons. These are the inner polygon, which is also called INNPOL, and the outer polygon, which is also called FIRPOL [6, 43, 44, 45] and which stands for *first polygon*. The outer polygon is a superset of $\mathcal{M}_S$, respectively $\mathcal{M}_C$, and contains all points in the $(s-1)$-dimensional AFS plane that represent nonnegative profiles. The precise mathematical definition of these outer polygons is as follows

$$\mathcal{F}_S = \left\{ x \in \mathbb{R}^{s-1} :\; V\begin{pmatrix}1\\ x\end{pmatrix} \geq 0 \right\}, \qquad \mathcal{F}_C = \left\{ y \in \mathbb{R}^{s-1} :\; U\Sigma\begin{pmatrix}1\\ y\end{pmatrix} \geq 0 \right\}. \tag{10}$$

The inner polygons are the convex hulls of the low-dimensional representations of the columns of D, respectively the rows of D. Let

$$a_i = \frac{(U\Sigma)^T(2:s,i)}{(U\Sigma)^T(1,i)}, \quad i = 1,\ldots,k, \quad \text{and} \quad b_j = \frac{(V(j,2:s))^T}{V(j,1)}, \quad j = 1,\ldots,n,$$

be the normalized vectors of expansion coefficients with respect to the left and right singular vectors of the columns and the rows of D, see also (6) and (7) in [46]. Then the inner polygons are the convex hulls of either the vectors a_i or the vectors b_j as

$$\mathcal{I}_S = \text{convhull}(\{a_i,\, i = 1,\ldots,k\}), \qquad \mathcal{I}_C = \text{convhull}(\{b_j,\, j = 1,\ldots,n\}). \tag{11}$$

The polygons $\mathcal{F}_C$, $\mathcal{I}_C$ for data set 2 are plotted on the left in figure 8, and the $\mathcal{F}_S$, $\mathcal{I}_S$ are shown on the right in the same figure. The polygons are drawn by black solid (inner polygon) and black broken lines (outer polygon). These plots additionally show the associated AFS-sets in transparent colors. Further, figure 9 exemplarily interprets the meaning of the outer polygon $\mathcal{F}_S$ for data set 2. Later in section 5.2 we discuss the MatLab program code to compute and to plot the inner and outer polygons. Their efficient computation exploits certain duality relations between $\mathcal{F}_S$ and $\mathcal{I}_C$ as well as $\mathcal{F}_C$ and $\mathcal{I}_S$, cf. section 3.6.

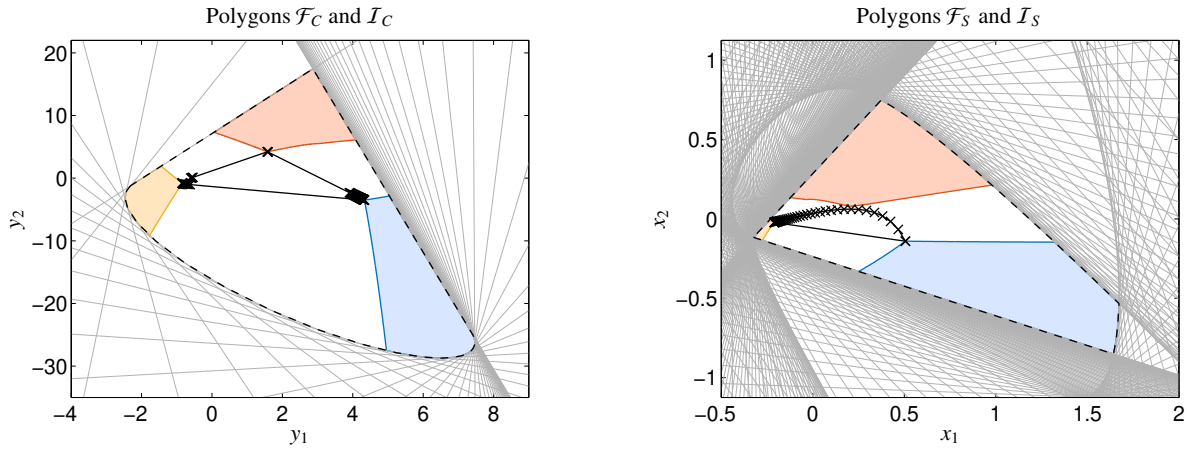

Figure 8: The polygons $\mathcal{F}_C$, $\mathcal{I}_C$ (left) and $\mathcal{F}_S$, $\mathcal{I}_S$ (right) as well as their construction for data set 2. The outer polygons (FIRPOL) $\mathcal{F}_C$ and $\mathcal{F}_S$ are plotted by black broken lines. The vectors $a(:, j)$ respectively $b(:, i)$ are marked by $\times$ symbols. The boundaries of their convex hulls are plotted by black solid lines. The gray lines limit the affine half-spaces $y_1(U\Sigma)_{j2} + y_2(U\Sigma)_{j3} \geq -(U\Sigma)_{j1}$, $j = 1, \ldots, k$, respectively $x_1 V_{i2} + x_2 V_{i3} \geq -V_{i1}$, $i = 1, \ldots, n$. The intersection of the y-half-spaces equals $\mathcal{F}_C$ and the intersection of the x-half-spaces forms $\mathcal{F}_S$. Additionally, the AFS-sets $\mathcal{M}_C$ and $\mathcal{M}_S$, each consisting of three isolated subsets, are plotted in transparent colors.

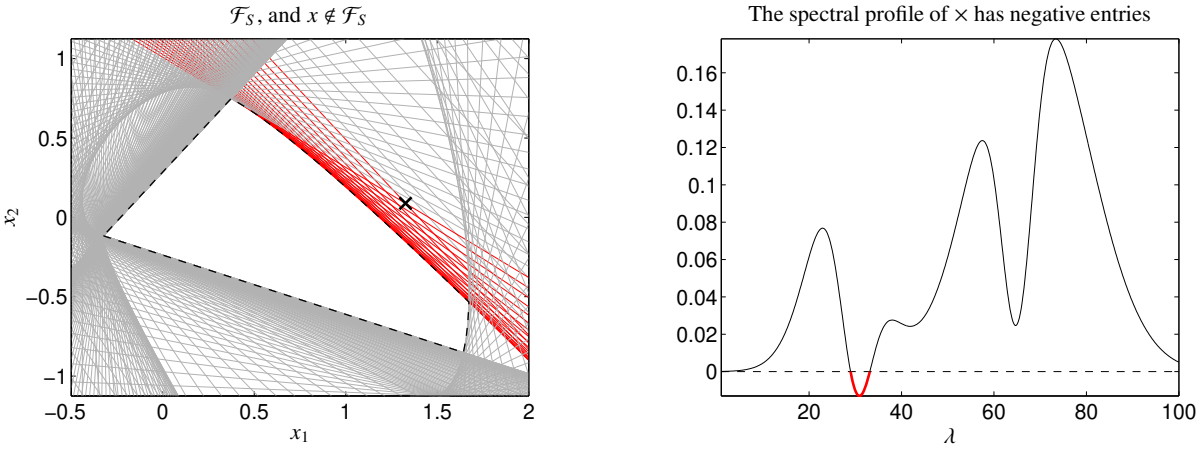

Figure 9: Data set 2: The figure illustrates (in the left subplot) the construction of $\mathcal{F}_S$ by the intersection of half-spaces that are implicitly defined by equation (10). The outer boundary of $\mathcal{F}_S$ is drawn by black broken lines. The gray lines are the limiting lines of various subspaces. The greater part of these lines does not contribute to the boundary of $\mathcal{F}_S$. An arbitrary point $x \notin \mathcal{F}_S$ is marked by a black cross. This point is not contained in some of the half-spaces. The boundary lines of these non-including spaces are drawn in red color. The right subplot shows the function which belongs to this point $x \notin \mathcal{F}_S$. This function has negative parts around $\lambda = 30$ (drawn red) and can therefore not be a feasible spectral profile.

3.3. On the history of the area of feasible solutions

In 1971 Lawton and Sylvestre presented not only the truncated SVD approach to (re)construct the factors C and S in [5]. They also analyzed the (rotational) ambiguity of the factorization problem for two-component systems. They presented their results in form of the so-called Lawton-Sylvestre plot of cones in the two-dimensional plane. There was no need to make use of the first-column-equal-1 scaling as introduced in (7) for the AFS approach. The Lawton-Sylvestre construction uses the permutation ambiguity and the single row representation with respect to T. The one-dimensional AFS for two-component systems can be extracted by intersecting the Lawton-Sylvestre plot by a line parallel to the ordinate and with the abscissa coordinate equal to 1.

The next developmental step goes back to Borgen and Kowalski in 1985 [6]. They extended Lawton-Sylvestre plots to three-component systems ($s = 3$) and introduced the seminal geometric interpretation of a feasible solution. Notable contributions also came from Sasaki, Kawata and Minami [47] in 1983 and Meister [48] in 1984. The only difference to the AFS-sets (8) and (9) is a somewhat different scaling. Borgen and Kowalski devised a geometric construction procedure [6] for the AFS that results in the so-called *Borgen-plots*.

The efforts to compute extremal profiles of the AFS by the MCR-BANDS software [31] are a different approach to estimate the degree of rotational ambiguity; see the work of Gemperline in 1999 [32] and Tauler in 2001 [21].

In 2005 Rajkó and István revived the Borgen plots by analytical investigations and by the application of tools from computational geometry [43]. In 2006 Vosough et al. presented the first generic numerical algorithm for the approximation of the AFS for ($s = 2$)-component systems [34], namely the grid-search method. This approach was analyzed in more detail by Abdollahi, Maeder and Tauler in 2009 [49]; therein the notion of the *area of feasible solution* was used for the first time, however in a slightly different manner.

In 2011 Golshan, Abdollahi and Maeder presented the triangle enclosure algorithm for approximating the boundary of the *area of feasible solutions* [8]. This method was the first fast (grid-search is comparatively slow) numerical algorithm to approximate the AFS. The method works stably for noisy data. An alternative computational method, namely the polygon inflation algorithm, was introduced in 2013 by Sawall, Neymeyr et. al. as an adaptive, stable and very fast numerical AFS-approximation method for three-component ($s = 3$) systems [9, 10]. In 2013 Golshan Maeder and Abdollahi extended the triangle enclosure algorithm for four-component systems ($s = 4$) by slicing the three-dimensional AFS and subsequent planar approximations by triangle enclosure [50]. In 2015 Jürß et al. presented an extension of the geometric Borgen plot construction [44, 45] to experimental spectral data sets including a low noise level and small negative data entries, which might result from a background subtraction. The ray casting algorithm for the computation of the AFS for systems with $s = 2, 3, 4, \ldots$ chemical components was introduced in 2017 [35]. The same research group has introduced the *dual Borgen plots* for the simultaneous computation of two dual AFS-sets $\mathcal{M}_S$ and $\mathcal{M}_C$ for three-component systems in extension of the geometric construction process. The dual Borgen plot construction is also possible for noisy data. However, the dual Borgen plot approach is not exclusively of geometric-constructive nature but includes in one step a polygon approximation by the polygon inflation procedure. See also [51, 52] for further reviews on the various developments.

3.4. Properties of the AFS

Next we summarize central properties of the AFS. The remaining part of section 3 has a mathematical character. All concepts (definitions) and theorems are accompanied by comments on their meaning and interpretation. Next restrict the presentation to properties of the AFS $\mathcal{M}_S$. Analogous properties hold for $\mathcal{M}_C$ due to the bilinear structure of the factorization. The point is that a transposition applied to $D = CS^T$ yields $D^T = SC^T$ and that this changes the positions of C and S. Next we need the term of an irreducible matrix.

Definition 3.1 (See also [41, 42])**.** *A symmetric $n \times n$-matrix H with $n \geq 2$ is called reducible if an $n \times n$ permutation matrix P exists so that*

$$PHP^T = \begin{pmatrix} H_{1,1} & 0 \\ 0 & H_{2,2} \end{pmatrix}.$$

Therein $H_{1,1}$ is an $m \times m$-matrix with $1 \leq m < n$. Otherwise, H is called irreducible.

For the following theory we have to assume the irreducibility of the two matrices $D^T D$ and DD^T for the spectral data matrix D. For experimental spectral data matrices D and for non-degenerated model data sets the irreducibility should always be satisfied. Otherwise the spectroscopic data can be separated to independent non-overlapping subsets, i.e. the reaction breaks down into two independent subreactions and the chemometric analysis can be applied separately to the subsets. Analogously, the set $\mathcal{M}_C$ can only be constructed if the matrix DD^T is irreducible.

Assumption 3.2. *D is a nonnegative matrix with* $\mathrm{rank}(D) = s$ *that has at least one factorization $D = CS^T$ with the nonnegative factors $C \in \mathbb{R}^{k \times s}$ and $S \in \mathbb{R}^{n \times s}$. Furthermore, $D^T D$ and DD^T are assumed to be irreducible matrices.*

As formally proved in theorem 2.2 of [10] the irreducibilty of $D^T D$ guarantees that the representation $T(:, 1) = (1, \ldots, 1)^T$ in (7) can be justified and that the AFS $\mathcal{M}_S$ can be constructed. The mathematical basis for this is the so-called Perron-Frobenius spectral theory of nonnegative matrices [53].

The assumption 3.2 is the basis for various important properties of the AFS. First, theorem 3.3 is about the geometric interpretation of a feasible solution in terms of a simplex which is located in the polygon $\mathcal{F}_S$ and which encloses $\mathcal{I}_S$. This theorem generalizes the classical construction principles of Borgen and Kowalski [6] of Borgen plots for two-component systems to systems with $s \geq 3$ chemical components.

Theorem 3.3 (Geometric condition on a feasible factorization; see [6, 43, 44, 46])**.** *On the assumption 3.2 a point $x \in \mathbb{R}^{s-1}$ belongs to $\mathcal{M}_S$ if and only if*

1. *x is contained in the outer polygon $\mathcal{F}_S$ and*
2. *further $s - 1$ points exist in $\mathcal{M}_S$ so that the simplex, which has these s points as vertices, encloses $\mathcal{I}_S$.*

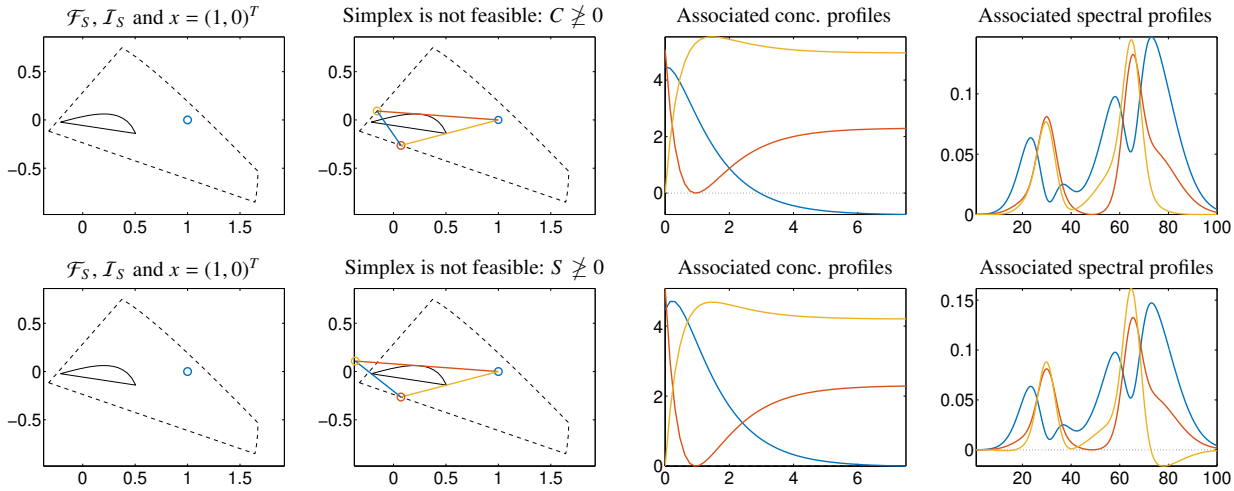

Figure 10: Check whether $x = (1,0)^T$ belongs to $\mathcal{M}_S$ for data set 2. Left column: x, $\mathcal{F}_S$ and $\mathcal{I}_S$ are plotted two times. There is no simplex with the vertex $x = (1,0)^T$ that is contained in $\mathcal{F}_S$ and includes $\mathcal{I}_S$. Two possible simplex-constructions are presented. Upper row: all vertices of the triangle are contained in $\mathcal{F}_S$ but one of its edges intersects $\mathcal{I}_S$. Hence one concentration profile has negative entries (upper row, blue concentration profile). All three spectra are nonnegative as all vertices of the triangle are located in $\mathcal{F}_S$. In the lower row, the triangle is constructed in a way that it tightly encloses the inner polygon but none of its edges are allowed to intersect the inner polygon. Then, necessarily one of its vertices is located outside the outer polygon. Therefore, the three associated concentration profiles are nonnegative, but one spectrum (drawn in ochre) has negative entries. Hence $x = (1,0)^T$ cannot be an element of $\mathcal{M}_S$.

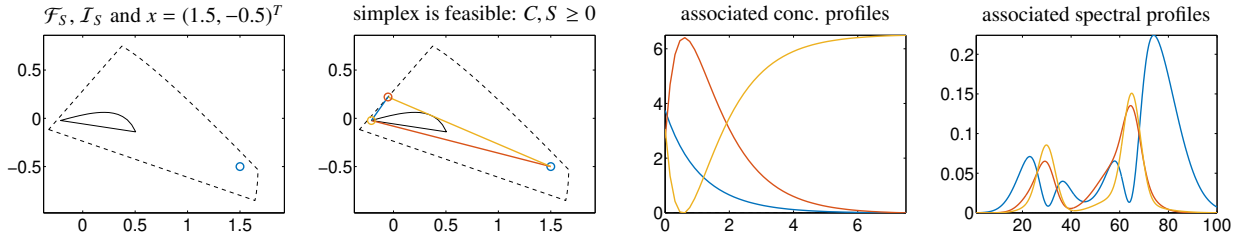

Figure 11: Check whether or not $x = (1.5,-0.5)^T$ belongs to $\mathcal{M}_S$ for data set 2. Left: x, $\mathcal{F}_S$ and $\mathcal{I}_S$. 2nd from left: Construction of a proper triangle that is contained in $\mathcal{F}_S$ and that encloses $\mathcal{I}_S$. As the construction succeeds, the associated concentration profiles (2nd from right) and spectral profiles (right) are all nonnegative. Hence $x = (1.5,-0.5)^T$ is an element of $\mathcal{M}_S$.

This theorem is explained for data set 2 in figures 10 and 11. In figure 10 we check whether $x = (1,0)^T$ belongs to $\mathcal{M}_S$. The first condition $x \in \mathcal{F}_S$ is fulfilled and it has still to be checked if two other points in $\mathcal{F}_S$ exist so that the associated triangle encloses $\mathcal{I}_S$. Such elements do not exist. In the upper line of plots in figure 10 a triangle with vertices in $\mathcal{F}_S$ is selected so that the factor S is nonnegative. However, one edge intersects the inner polygon $\mathcal{I}_S$ and so the associated concentration profile has some negative parts. In the lower line of plots in figure 10 a triangle which tightly encloses $\mathcal{I}_S$ is constructed. Hence C is nonnegative. However, one vertex is outside $\mathcal{F}_S$, so the associated spectral profile has negative entries. In figure 11 we check whether $x = (1.5,-0.5)^T$ belongs to $\mathcal{M}_S$. Since $x \in \mathcal{F}_S$ the first condition is fulfilled. Furthermore, two other points in $\mathcal{F}_S$ exist so that the triangle (with these three points as vertices) encloses $\mathcal{I}_S$. Hence, it holds $(1.5,-0.5)^T \in \mathcal{M}_S$. These two additional vertices of the triangle are not uniquely determined.

The next theorem summarizes some simple geometric properties of the AFS $\mathcal{M}_S$.

Theorem 3.4 (Summary of properties from [10, 35, 54])**.** *On the assumption 3.2 it holds that:*

1. *The outer polygon $\mathcal{F}_S$, called FIRPOL in [6], is a bounded set. Thus its subset $\mathcal{M}_S$ is also a bounded set.*
2. *The origin, i.e. the null vector, is an element of the outer polygon $\mathcal{F}_S$ but it is not contained in $\mathcal{M}_S$.*
3. *The intersection of $\mathcal{M}_S$ with a ray starting at the origin is either empty or a (gap-free) interval.*
4. *The set $\mathcal{M}_S$ is a closed set (in the topological sense of closed sets, namely that all boundary points belong to the set).*

All these geometric properties can be simply checked by the reader for any AFS, e.g. for the AFS sets shown in figure 13. Some immediate consequences of this theorem for general AFS are as follows: If $\mathcal{M}_S$ consists of more than one (isolated) subset, then these subsets cannot contain holes. In the other case, namely that the AFS continuously surrounds the origin, the origin with its neighborhood is the only hole in the AFS.

3.5. The case of two-component systems

For two-component systems ($s = 2$) it has become standard practice [34, 49, 55] to work with a matrix T in the form of

$$T = \begin{pmatrix} 1 & t_{12} \\ t_{21} & 1 \end{pmatrix}, \tag{12}$$

instead of using T by (7). The AFS representations according to (12) and (7) differ in terms of their scaling. In each of these two cases the AFS is one-dimensional and consists of two separated finite intervals of real numbers.

In order to analyze the relations between these AFS-representations let $D = CS^T$ be a nonnegative factorization which allows us to form the matrix $\tilde{T} = S^T V$. With the theory underlying theorems 3.3 and 3.4 one can show that $\tilde{T}_{11} > 0$, $\tilde{T}_{21} > 0$ as well as $\tilde{T}_{12} \cdot \tilde{T}_{22} < 0$. After a proper (re)scaling and with the approach from (12) we have $t_{12} < 0$ and $t_{21} > 0$. Hence, if the one-dimensional AFS $\mathcal{M}_S$ equals the union of two intervals in the form $\tilde{\mathcal{M}}_S = \tilde{I}_- \cup \tilde{I}_+$ with respect to T by (12) and if $\mathcal{M}_S = I_- \cup I_+$ is the AFS with respect to (7) with $\tilde{I}_-, I_- \subset \{x \in \mathbb{R} : x < 0\}$ and $\tilde{I}_+, I_+ \subset \{x \in \mathbb{R} : x > 0\}$, then it holds that $\tilde{I}_- = I_-$ and $\tilde{I}_+ = \{x^{-1} : x \in I_+\}$.

3.6. Duality relations between outer polygons and inner polygons

This section discusses relations between the AFS for the spectral factor S and the AFS for the concentration factor C. Due to the equation $D = CS^T$ the factors C and S are closely coupled and so are various geometric properties of these AFS sets and their construction principles. Unfortunately, the duality theory is somewhat complicated and we cannot introduce and explain the geometric and mathematical background here. Instead, we must refer to the numerous references.

The geometric interpretation by theorem 3.3 and the geometric construction of the AFS-sets are based on the polygons $\mathcal{F}_S$ and $\mathcal{I}_S$ respectively $\mathcal{F}_C$ and $\mathcal{I}_C$. For ease of presentation we prefer to talk of polygons. However, in the general case $s > 3$, the four sets $\mathcal{F}_S, \mathcal{I}_S, \mathcal{F}_C$ and $\mathcal{I}_C$ are polyhedra. These four polygons (polyhedra) are convex and bounded subsets in $\mathbb{R}^{s-1}$. A direct computation of the outer polygons $\mathcal{F}_S$ and $\mathcal{F}_C$ can be computationally expensive. Further, a direct computation of $\mathcal{I}_S$ and $\mathcal{I}_C$ can be hard for noisy data. These two problems can be solved by means of duality relations between the geometric objects.

The *duality principle* for nonnegative matrix factorizations in the low-dimensional AFS representation is described in [56, 57, 46, 58]. See also the closely related complementarity theory in [59, 60, 61, 62]. Duality-based algorithms enable a fast computation of the polygons $\mathcal{F}_S$ and $\mathcal{F}_C$ for (noise-free) model data and also to approximate the polygons $\mathcal{I}_S$ and $\mathcal{I}_C$ for noisy data. In theorem 3.5 we describe duality relations between the columns of C and S in the AFS-space representation.

Theorem 3.5 (Duality relations in the AFS-space [57, 59, 46])**.** *Let $D = CS^T$ be a nonnegative matrix factorization according to* (6) *so that $T = S^T V$. Further, let the row vectors $x^{(i)}$ be the representatives of the columns of S, and let the column vectors $y^{(j)}$ be the analogous representatives of the columns of C, i.e.*

$$x^{(i)} = \frac{T(i, 2:s)}{T_{i1}}, \quad i = 1, \ldots, s, \qquad\qquad y^{(j)} = \frac{(T^{-1})(2:s, j)}{(T^{-1})_{1j}}, \quad j = 1, \ldots, s.$$

Then it holds for $i \neq j$ that

$$\sum_{\ell=1}^{s} x_\ell^{(i)} y_\ell^{(j)} = -1.$$

The following definition prepares the ground for explaining the duality relations between the polygons in the AFS-space.

Definition 3.6. *A nonzero vector $x \in \mathbb{R}^{s-1}$ and an affine hyperplane $E = \{y \in \mathbb{R}^{s-1} : y^T x_E = -1\}$ are called dual if $x_E = x$.*

In words, definition 3.6 says that a nonzero vector x determines its dual affine hyperplane and that an affine hyperplane (not including the origin) determines its dual generating vector x. The next theorem 3.7 describes duality relations between $\mathcal{F}_S$ and $\mathcal{I}_C$ and the similar relations between $\mathcal{F}_C$ and $\mathcal{I}_S$.

Theorem 3.7 (Duality between inner and outer polygons [56, 57, 46])**.** *The vertices (respectively faces) of $\mathcal{F}_S$ are dual to the faces (respectively vertices) of $\mathcal{I}_C$. Analogically, the vertices (resp. faces) of $\mathcal{F}_C$ are dual to the faces (resp. vertices) of $\mathcal{I}_S$.*

With the help of theorem 3.7 the inner polygons and the outer polygons can easily be computed as follows:

- **For noise free data** one can compute the inner polygon $\mathcal{I}_S$ as the convex hull of the a_i, $i = 1, \ldots, k$, and the inner polygon $\mathcal{I}_C$ as the convex hull of the b_j, $j = 1, \ldots, n$. (The numerical computing environment MatLab provides the routine *convhull* for these computations.) Then the vertices of $\mathcal{F}_C$ are the dual vectors of the edges of $\mathcal{I}_S$. Analogously, the vertices of $\mathcal{F}_S$ are the dual points of the edges of $\mathcal{I}_C$.
- **For noisy data** we compute approximations to the outer polygons $\mathcal{F}_S$ and $\mathcal{F}_C$ for example by the (inverse) polygon inflation procedure, see section 4.6 or [10]. Then the vertices of $\mathcal{I}_C$ are determined as the dual vectors of the edges of $\mathcal{F}_S$ and the vertices of $\mathcal{I}_S$ as the dual vectors of the edges of $\mathcal{F}_C$. See [58] for details on this procedure.

Theorem 3.7 also applies to any dimension $s \geq 3$. Then we consider inner and outer polyhedra (instead of polygons) and faces of the polyhedra instead of edges.

3.7. Approximations of the inner/outer polygons for noisy data

The strict definition of the AFS-sets does not allow any deviation from the nonnegativity $C, S \geq 0$. However, for experimental noisy and distorted spectral data one should accept small negative entries in D and also in its factors C and S. Without accepting such entries, the AFS-sets might be empty. We recommend the following lower bounds on the relative size of the acceptable negative entries

$$\frac{C_{i\ell}}{\max|C(:,\ell)|} \geq -\varepsilon_C, \quad i = 1, \ldots, k, \qquad \text{and} \qquad \frac{S_{j\ell}}{\max|S(:,\ell)|} \geq -\varepsilon_S, \quad j = 1, \ldots, n, \tag{13}$$

for $\ell = 1, \ldots, s$ with small control parameters $\varepsilon_C, \varepsilon_S \geq 0$. The polygon inflation algorithm [9, 10] and the ray casting method [35] work with these bounds. A different approach is to set all negative entries equal to zero and to use the ssq-function (the Frobenius norm) as a measure of the distance of the spectral data to its reconstruction [34, 49, 8, 50].

If the constraints (13) are used, then the polygons $\mathcal{F}_S$, $\mathcal{F}_C$, $\mathcal{I}_S$ and $\mathcal{I}_C$ differ slightly from their form for strict nonnegativity constraints. In order to compute the AFS for noisy and experimental data, the publication [58] suggests the approach to compute first approximations of $\mathcal{F}_S$ and $\mathcal{F}_C$ with respect to (13) and then to construct $\mathcal{I}_S$ and $\mathcal{I}_C$ by using the underlying duality. Figure 12 shows $\mathcal{F}_S$, $\mathcal{F}_C$, $\mathcal{I}_S$ and $\mathcal{I}_C$ for data set 3 with respect to (13) and $\varepsilon_C = \varepsilon_S = 6 \cdot 10^{-3}$. The a_i, $i = 1, \ldots, k$ respectively b_j, $j = 1, \ldots, n$, are also drawn together with the affine hyperplanes that define $\mathcal{F}_S$ and $\mathcal{F}_C$.

3.8. Orientation of the AFS plots

The AFS-sets decisively depend on the choice of the singular vectors. Although the singular values of D are unique, the singular vectors are not uniquely determined even if all singular values are simple (i.e. have the multiplicity 1). Here we do not consider the case of multiple singular values with their high degree of freedom in the determination of the singular vectors. For simple singular values only the orientation of the singular vectors (namely positive or negative sign) remains undetermined. However, the first left- and right singular vectors are sign-constant due to the Perron-Frobenius theory (i.e. either nonpositive or nonnegative) and we assume $U(:,1) > 0$ and $V(:,1) > 0$.

The sign-indeterminacy of all other singular vectors results in an $\pm$-indeterminacy of the orientation of the axes in the AFS plots. In order to cure this problem and to improve the comparability of AFS plots we set up the following convention on the right singular vectors $V(:,j)$ (which also determines the signs of the left singular vectors)

$$\max V(:,i) \geq -\min V(:,i), \qquad i = 2, \ldots, s. \tag{14}$$

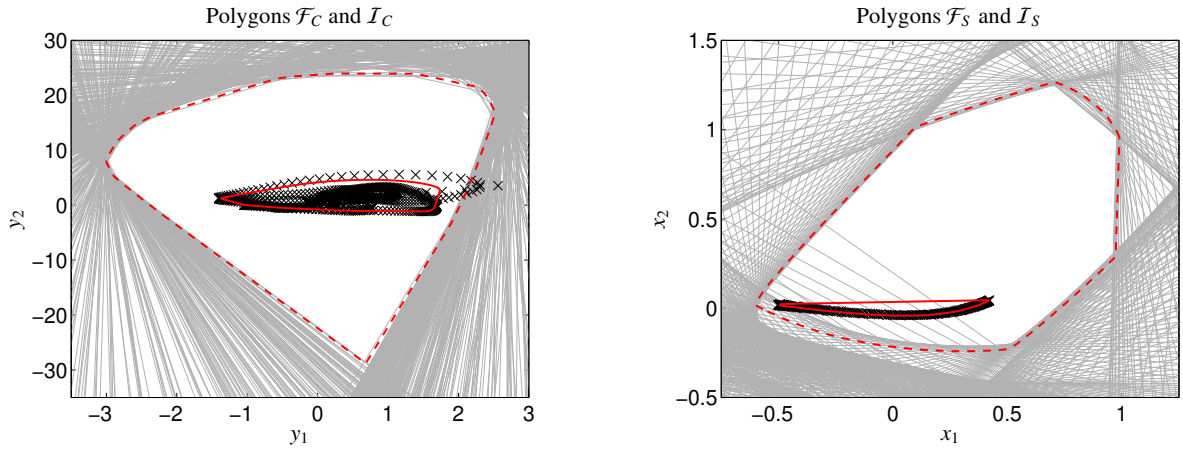

Figure 12: For data set 3 the polygons $\mathcal{F}_C, \mathcal{I}_C$ (left) and $\mathcal{F}_S, \mathcal{I}_S$ (right) are drawn with respect to the weakened nonnegativity constraint (13). The gray lines limit the affine half-spaces $y_1(U\Sigma)_{j2} + y_2(U\Sigma)_{j3} \geq -(U\Sigma)_{j1}$, $j = 1, \ldots, k$, respectively $x_1 V_{i2} + x_2 V_{i3} \geq -V_{i1}$, $i = 1, \ldots, n$. Hence these straight lines serve to construct the boundaries of $\mathcal{F}_C$ and $\mathcal{F}_S$ with respect to the strict/non-weakened nonnegativity constraint. If small negative matrix entries in the sense of (13) are accepted, then $\mathcal{F}_S$ and $\mathcal{F}_C$ increase in size. We use $\varepsilon_C = \varepsilon_S = 6 \cdot 10^{-3}$. The resulting boundaries of $\mathcal{F}_S$ and $\mathcal{F}_C$ are drawn as red broken lines. The data representing vectors a_i and b_j are marked by black $\times$ symbols. In the left plot, some of the a_i are located outside $\mathcal{F}_C$ (which would be impossible for strictly nonnegative data and which would make an AFS construction impossible). This problem can be cured by exploiting the duality relations. First, $\mathcal{F}_S$ determines the dual polygon $\mathcal{I}_C$ (red solid line in the left subplot). Second, $\mathcal{F}_C$ determines the dual polygon $\mathcal{I}_S$ (red solid line in the right subplot). Without these modified polygon constructions, the two AFS-sets would be empty since no triangle exists in $\mathcal{F}_S$ (respectively $\mathcal{F}_C$) that encloses $\mathcal{I}_S$ (respectively $\mathcal{I}_C$), see theorem 3.3.

In words, equation (14) means that the component of the largest modulus is assumed to be positive. Practically, changing the sign of the i_0th singular vector means that the associated AFS plot has a flipped x_{i_0-1}-axis. The possible axis orientations of a two-dimensional AFS for a chemical three-component system are shown in figure 13 for data set 3.

4. Methods for the construction and numerical approximation of the AFS

This section reviews some of the techniques for the construction or numerical approximation of the AFS in more detail, cf. the list of methods introduced in section 3.3. First, we discuss a numerical test on the feasibility, namely whether a certain point x belongs to the AFS.

4.1. Numerical feasibility test

A central ingredient of any numerical AFS-approximation method is the numerical test in order to decide if a certain x is feasible. Next, we describe such a test applicable to model data and to noisy experimental data. This approach uses the weakened nonnegativity constraints (13) that tolerate small negative entries in C and S. The starting point is the objective function

$$f(x, W) = \sum_{\ell=1}^{s} \sum_{i=1}^{k} \left(\min\left(0, \frac{C_{i\ell}}{\max C(:,\ell)} + \varepsilon_C\right)\right)^2 + \sum_{\ell=1}^{s} \sum_{j=1}^{n} \left(\min\left(0, \frac{S_{j\ell}}{\max S(:,\ell)} + \varepsilon_S\right)\right)^2 + \left\| I_s - T^+ T \right\|_F^2 \qquad (15)$$

with T by (7) and C and S by (6). A vector $x \in \mathbb{R}^{s-1}$ is accepted as an element of the AFS if and only if

$$F(x) = \min_{W \in \mathbb{R}^{s-1 \times s-1}} f(x, W) < \varepsilon_f \qquad (16)$$

with a small control parameter ε_f. A typical choice is $\varepsilon_f = 10^{-12}$. For example, this test is used in [9, 10, 35]. It is also used for all numerical AFS approximation methods in this chapter.

Each evaluation of $F(x)$ requires the solution of a nonlinear least-squares problem with $(s-1)^2$ variables and a residual vector with $s(k + n + s)$ components. We use the powerful adaptive nonlinear least-squares program code NL2SOL [63, 64] in our software. In order to make the implementation of the feasibility test efficient, we apply a first-stage rapid test (whether x is contained in the outer polygon). If this test is not successfully passed, then $x \notin \mathcal{M}_S$. Otherwise, if x is located in the outer polygon, then the computationally expensive test based on (15) and (16) is applied to test the AFS membership [10].

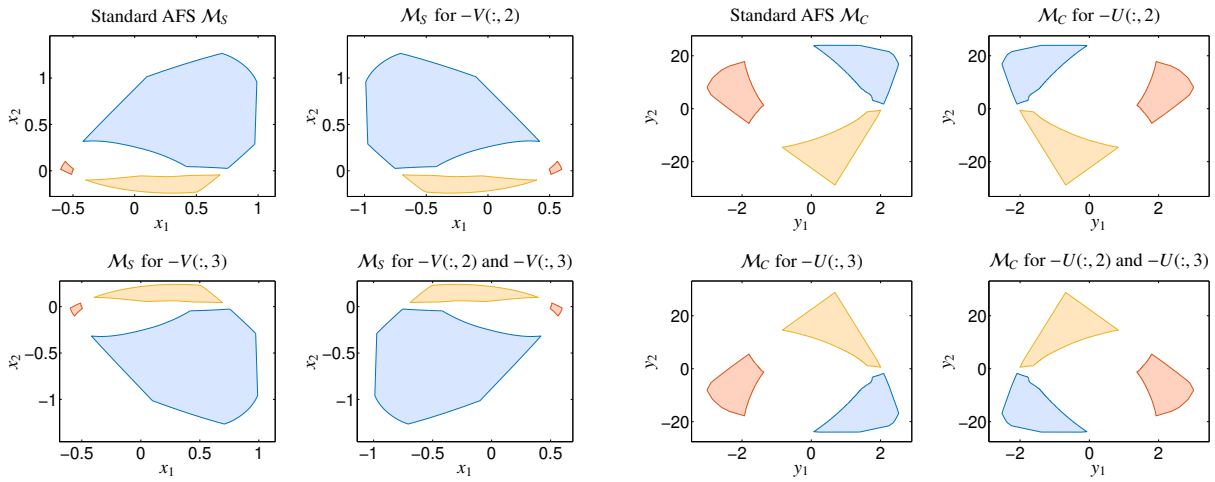

Figure 13: Possible axis orientations of a two-dimensional AFS for the chemical three-component systems underlying data set 3. The standard axis orientations of $\mathcal{M}_S$ and $\mathcal{M}_C$ use the sign selection rule specified in (14). The three other AFS-plots each for $\mathcal{M}_S$ and $\mathcal{M}_C$ refer to possible sign changes of the second and third singular vector and correspond to reflections along the different axes.

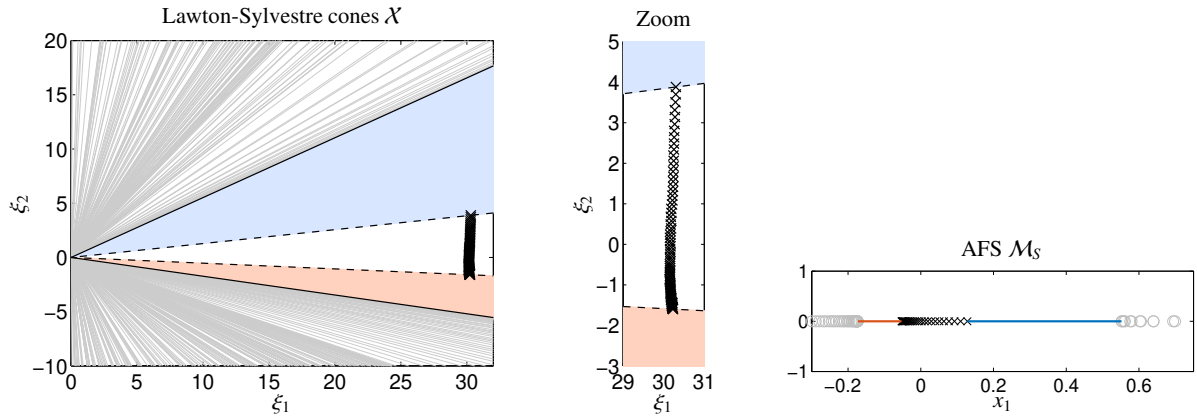

Figure 14: The Lawton-Sylvestre plot (left) for data set 1. The gray lines are the boundaries of the affine half-spaces $\xi_1 \geq -\xi_2 V_{i2}/V_{i1}$, $i = 1\ldots,n$, (only every fifth line is plotted). The black lines define the "outer" boundaries with respect to the nonnegativity constraint. The crosses are the spectral data representatives $(\sigma_1 U_{j1}, \sigma_2 U_{j2})$, $j = 1,\ldots,k$. The broken black lines are the "inner" boundaries of the two cones of feasible points ξ that are drawn as blue and red areas. The two feasible cones are plotted in light blue and light red. The associated AFS $\mathcal{M}_S$ (right plot) is the intersection of the two cones with the line with the fixed coordinate $\xi_1 = 1$.

4.2. Lawton-Sylvestre plots for $s = 2$ and $s = 3$

The Lawton-Sylvestre plots [5] can be computed for two- and three-component systems and represent graphical representations in two or three dimensions. These plots have the same information content as the AFS-plots, but do not make use of the leading-1 normalization as introduced in (7). Thus Lawton-Sylvestre plots have a dimension that is greater by 1 compared to AFS plots for the same system. Feasible regions are formed by the points (ξ_1, ξ_2) so that

$$S(:,1)^T = \xi_1 V(:,1)^T + \xi_2 V(:,2)^T$$

is a possible profile of a nonnegative factorization of the spectral mixture data matrix D. If the SVD is selected in a way that $V(:,1) > 0$, then theorem 2.2 in [10] predicts that $\xi_1 > 0$. One can show that the set of feasible (ξ_1, ξ_2) has the form

$$\mathcal{X} = \left\{(\xi_1,\xi_2):\ \xi_1 > 0 \text{ and } \left(-\min_{i:\, V_{i2}>0} \frac{V_{i1}}{V_{i2}} \leq \frac{\xi_2}{\xi_1} \leq \min_i \frac{\sigma_2 U_{i2}}{\sigma_1 U_{i1}} \text{ or } \max_i \frac{\sigma_2 U_{i2}}{\sigma_1 U_{i1}} \leq \frac{\xi_2}{\xi_1} \leq -\max_{i:\, V_{i2}<0} \frac{V_{i1}}{V_{i2}}\right)\right\}.$$

This set consists of two semi-infinite cones. The apex of each cone is the origin (the null vector). However, the origin does not belong to the cone. Figure 14 shows the spectral factor cone $\mathcal{X}$ for data set 1. The intersection of $\mathcal{X}$ with the straight line with the fixed coordinate $\xi_1 = 1$ results in the one-dimensional AFS $\mathcal{M}_S$ as considered in this paper.

Lawton-Sylvestre plots can also be computed for $s \geq 3$. However, their visual presentation in higher dimensions is not very informative. For $s = 3$ the set $\mathcal{X}$ can consist of three-dimensional semi-infinite cones. Once again, each

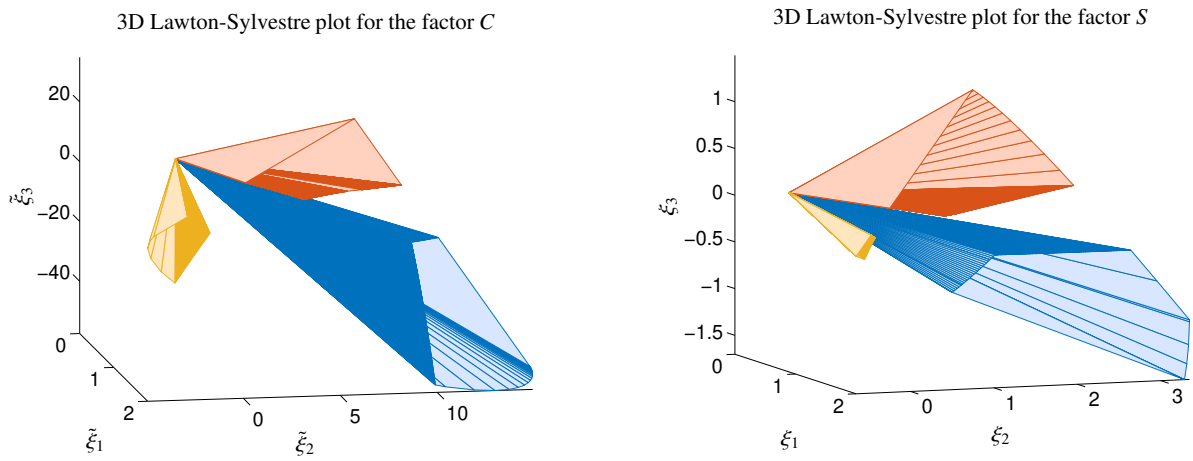

Figure 15: Three-dimensional Lawton-Sylvestre plots for data set 2 with three chemical components.

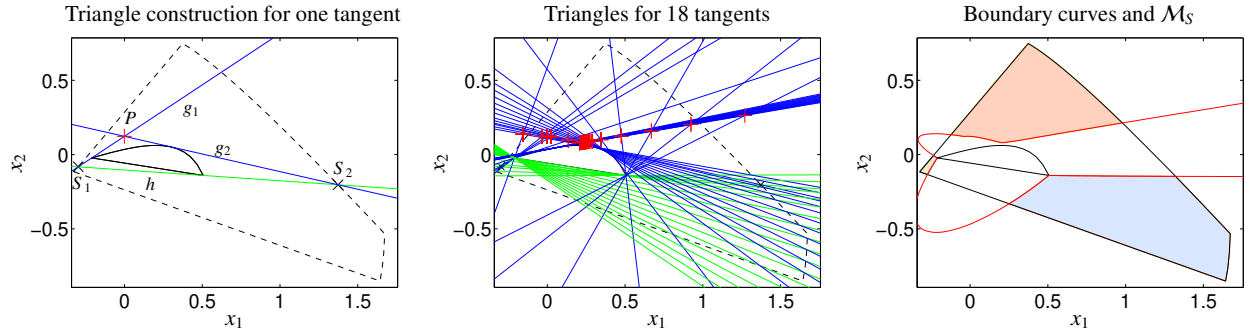

Figure 16: Geometric construction of the inner boundary curve of the spectral AFS for data set 2. Left: One tangent h (green line) of the inner polygon is shown which touches $\mathcal{I}_S$ (black solid line). The two points of intersection S_1 and S_2 (×) of h with the boundary of $\mathcal{F}_S$ (black broken line) are marked. Starting from these two points, the two lines g_1 and g_2 (blue) are computed as tangents of $\mathcal{I}_S$. The intersection of g_1 and g_2 (+) is P. Since P is in $\mathcal{F}_S$, this point is an inner boundary point of $\mathcal{M}_S$. Center: The same construction for 18 different tangents results in 18 boundary points. Right: The curve through these boundary points (red solid line, for 1000 tangents) covers the inner boundary of $\mathcal{M}_S$. In combination with the boundary curve of the outer polygon $\mathcal{F}_S$ the spectral AFS $\mathcal{M}_S$ (colored areas) is shown. For this data set the AFS consists of three isolated subsets.

cone has its apex in the origin (the null vector), but the apex does not belong to the cone. Figure 15 shows typical 3-dimensional Lawton-Sylvestre plots, namely for data set 2.

4.3. Borgen plots by geometric construction

The geometric construction of Borgen plots for three-component systems is an elegant and fast approach to construct the AFS [6, 43, 44, 46]. Essentially, the geometric construction yields the "inner" parts of the boundary of the AFS. The remaining parts of the boundary, namely the "outer" boundary belongs to the boundary of the outer polygon.

4.3.1. Classical Borgen plots

The theoretical background for the construction of Borgen plots is theorem 3.3. For a three-component system a certain point in the AFS-plane belongs to its inner boundary if it is contained in the outer polygon $\mathcal{F}_S$ and if two additional points on the boundary of $\mathcal{F}_S$ exist so that the triangle spanned by these three vertices tightly encloses $\mathcal{I}_S$. For details see lemma 4.3 in [46]. Thus the inner boundary can be constructed either by an analytic curve construction [43] or by a triangle rotation process. In the latter process a tangent rotates around $\mathcal{I}_S$ and for every tangent h a triangle is computed with one edge being h. The triangle tightly encloses $\mathcal{I}_S$ and has two vertices on the boundary of $\mathcal{F}_S$. If the third vertex is in $\mathcal{F}_S$, then this vertex is an inner boundary point, see lemma 4.3 of [46]. The set of all these points forms the inner boundary. Figure 16 illustrates the geometric construction of the inner boundary for the spectral AFS and data set 2.

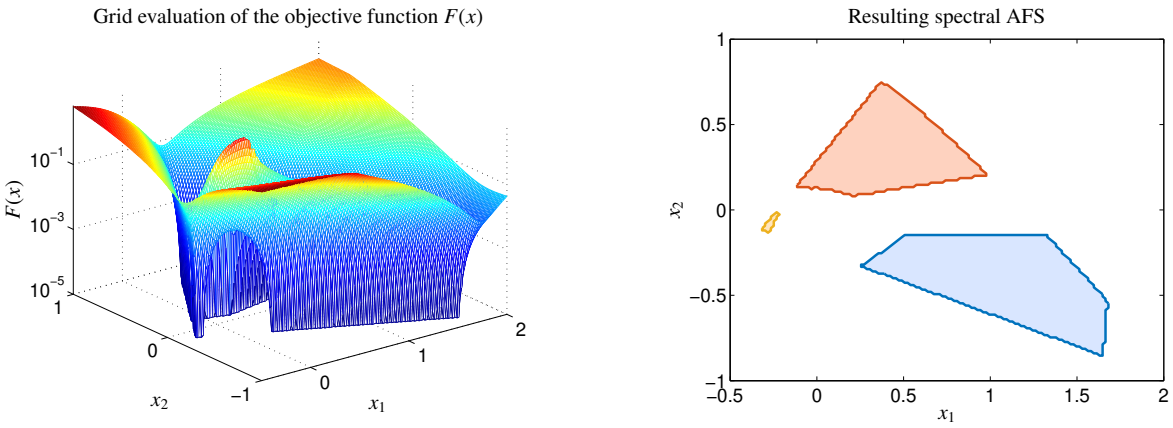

Figure 17: Application of the grid-search method in order to approximate $\mathcal{M}_S$ for data set 2. Left: the objective function $F(x) = \min_S f(x, S)$ by (16) is evaluated on the rectangle $[-0.5,\ 2] \times [-1,\ 1]$ discretized on a mesh with 150×151 nodes. Right: the corresponding AFS $\mathcal{M}_S$ for $\varepsilon_f = 10^{-5}$.

4.3.2. Generalized Borgen plots

Generalized Borgen plots extend the geometric AFS construction from the noise-free (model data) case to the noise-loaded (experimental data) case [44, 45]. In this approach convex combinations with nonnegative expansion coefficients are substituted by affine combinations that use triangles whose vertices can slightly leave $\mathcal{F}_S$. Thus small negative entries in S are accepted. Furthermore, the restriction that the triangles have to enclose $\mathcal{I}_S$ is slightly broken in order to accept small negative entries also in the factor C. This approach does not use the constraint (13) on the *relative* size of negative entries but works with lower bounds on the negative entries (in a non-relative way).

4.3.3. Dual Borgen plots

The recent approach of dual Borgen plots extends the construction of a single Borgen plot to the simultaneous constructions of the two Borgen plots for the spectral factor and the concentration factor. The simultaneous construction using duality relations between the two AFS-sets speeds up the process considerably. The general algorithm is introduced in [46]. In [58] an algorithmic extension for noisy data is explained. The method accepts negative entries in the sense of (13) and works in a stable way even for noisy data. Compared to the classical Borgen plots not only one triangle is constructed per tangent but three of them. The two other triangles lead (via the duality relations) to the inner boundary points of $\mathcal{M}_C$. For noisy data, first approximations of $\mathcal{F}_S$, $\mathcal{F}_C$, $\mathcal{I}_S$ and $\mathcal{I}_C$ are computed (cf. figure 12) and then the geometric constructions are carried out for these polygon approximations.

4.4. Grid-search for $s \geq 2$

The grid-search method is a brute-force numerical algorithm in order to compute the AFS-sets [34, 49]. The method works, in principle, for any $s \geq 2$. It can be applied to noisy data. First a grid in the $\mathbb{R}^{s-1}$ is defined and then the numerical feasibility test by means of the function $F(x)$ by (16) is executed for all nodes of the grid. The accuracy of the approximation depends on the grid width of the discretization. Figure 17 illustrates the grid-search process for $h_{x_1} = 1.68 \cdot 10^{-2}$, $h_{x_2} = 1.33 \cdot 10^{-2}$ and $\varepsilon_f = 10^{-5}$ for the spectral AFS and data set 2. In this context, we also refer to the particle swarm optimization as presented in [65].

4.5. Triangle enclosure for $s = 3$

The triangle enclosure algorithm is a method for the approximation of the AFS for three-component systems [8]. It works in a stable way also for noisy data. The idea is to compute a chain of equilateral triangles that covers the boundary of the AFS. Each triangle has at least one vertex in $\mathcal{M}_S$ and at least one vertex outside $\mathcal{M}_S$. Two consecutive triangles share a joint edge. If the AFS consists of three separated subsets, then the algorithm is applied to each of the subsets. If the AFS is a topologically connected set containing a hole, then two triangle chains are generated, one for the inner and one for the outer boundary. The accuracy of the boundary approximation depends on the side length.

The triangle enclosure algorithm works as follows: At an initialization phase a first feasible point is assumed to be given in the subset of the AFS that is to be approximated. Such a feasible point can be computed by an NMF

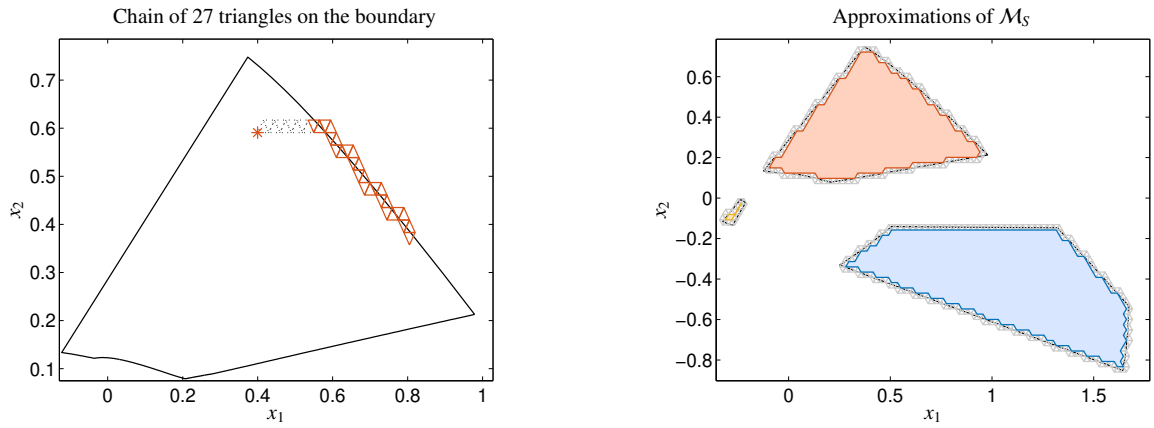

Figure 18: Approximation of $\mathcal{M}_S$ by the triangle enclosure algorithm for data set 2. The edge length of the equilateral triangles is 0.03. Left: the initialization phase starts with an interior feasible point (✱) and builds a straight chain of triangles (broken black lines). A chain of 27 triangles covering a part of the boundary of the AFS is drawn in red. Right: finally all feasible vertices of the vertices in the triangle chain are piecewise linearly connected, which forms the approximation of the boundary of $\mathcal{M}_S$.

of the given spectral mixture data matrix. This feasible point is taken as a first vertex of an arbitrary equilateral triangle. Its side length is equal to a given pre-defined small number that finally determines the quality of the boundary approximation. If none of its vertices leaves the AFS, then the triangle is reflected along one of its edges to form a new triangle. If none of the vertices leaves the AFS, then the reflection process is continued in a way that forms a straight chain of triangles, see the left plot in figure 18. If a first triangle has at least one vertex in the AFS and at least one vertex outside the AFS, then the initialization phase comes to an end and the iteration phase starts. Let Δ be an equilateral triangle that crosses the boundary of $\mathcal{M}_S$ in the sense that one vertex (P_1) is feasible and one vertex (P_2) is not feasible. The third vertex can be feasible or not feasible. Next the current triangle is reflected along the edge from P_1 to P_2. This results in a new triangle Δ' with a new, uniquely determined vertex P'_3. Again, the new triangle is reflected either along the edge from P'_3 to P_1 or from P'_3 to P_2 (the algorithm uses for this reflection the edge that crosses the boundary of the AFS). The iteration stops if the triangle chain closes. The process is illustrated in figure 18 for data set 2 and with the side length equal to 0.03.

In [50] an extension for the sliced approximation of the AFS-sets for four-component systems ($s = 4$) is introduced. Therein the slices are generated along the x_3 coordinate. For each slice a two-dimensional AFS is computed by the triangle enclosure algorithm.

4.6. Polygon inflation algorithms for $s = 3$

The polygon inflation algorithm [9] and the inverse polygon inflation procedure [10] are numerical methods for the approximation of the AFS boundary (respectively for each of its subsets) by an adaptively refined polygon in the case $s = 3$. For further references see [51, 52, 66]. The method facilitates a very accurate boundary approximation. It is fast and stable even for noisy data.

The idea of polygon inflation is to approximate the boundary curve of each AFS subset by a sequence of adaptively refined, inflating polygons. The vertices of these polygons are (with a high precision) located on the boundary of $\mathcal{M}_S$. The starting point is a single feasible solution (e.g. from an initial NMF). A triangle is constructed around this initial point so that its three vertices are all located on the boundary of the AFS. This triangle is a first, coarse polygon approximation of the AFS. Then the polygon is iteratively refined by edge subdivisions. To this end, the algorithm computes a new vertex as the point of intersection of the perpendicular bisector of a given edge and the boundary of the AFS (subset). If the new vertex is outside the old polygon, then the current step is an *inflation*. In some cases, the new vertex is located in the current polygon. The latter case amounts to a *deflation*. If a new edge is computed, then a correction-distance parameter is stored in a vector Δ. Based on all correction-distances of the current polygon, the next edge is selected in a way that a further edge refinement is executed in regions of the boundary where the past refinement step has provided a large correction-distance. This is the idea of the adaptive process. The iteration continues until all correction-distances fall below a critical value δ. This defines the stopping criterion. With certain modifications the method can also compute punctiform or line-shaped subsets of the AFS. Figure 19 shows a flow chart of the polygon inflation algorithm. The progress of the approximation process of $\mathcal{M}_C$ by polygon inflation is presented in figure 20 for data set 2.

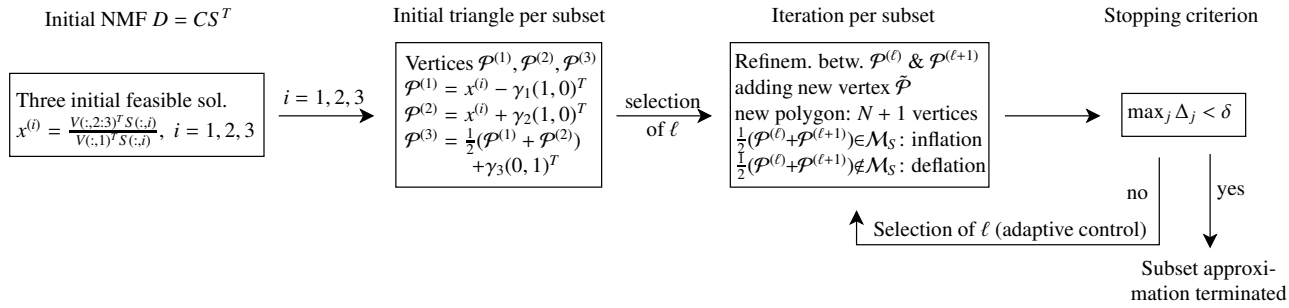

Figure 19: Flow chart of the polygon inflation algorithm.

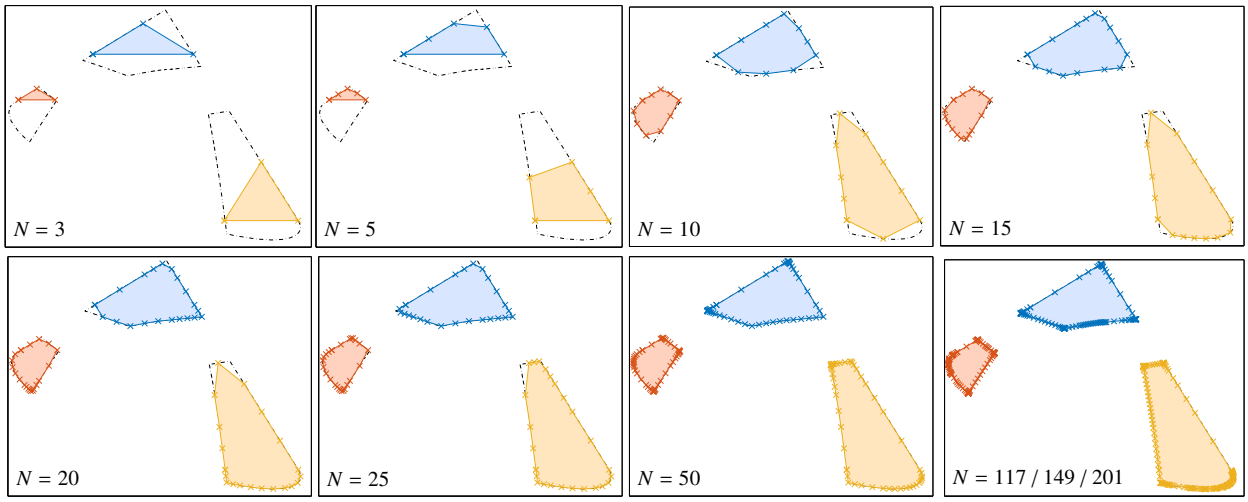

Figure 20: Progress of the adaptive polygon approximations by means of the polygon inflation algorithm in order to compute $\mathcal{M}_C$ for data set 2. The total number of current vertices for the three subsets is N. The final boundaries of the AFS are drawn by black dash-dotted lines.

The inverse polygon inflation algorithm works as follows: First the polygon inflation procedure is used in order to compute the outer polygon $\mathcal{F}_S$. Then a second polygon $\mathcal{M}_S^*$ with

$$\mathcal{M}_S^* = \left\{ x \in \mathbb{R}^2 : \ \exists W \in \mathbb{R}^{2\times 2} \text{ by (7) with } \operatorname{rank}(T) = 3,\ U\Sigma T^{-1} \geq 0,\ (TV^T)(2:3,:) \geq 0 \right\} \tag{17}$$

is determined, which serves to determine the inner boundary of $\mathcal{M}_S$. For details see [10]. Finally, the desired AFS is constructed as the intersection $\mathcal{M}_S = \mathcal{F}_S \cap \mathcal{M}_S^*$. The inverse polygon inflation method is the suitable method if the AFS consists of a contiguous area containing a hole (around the origin). Moreover, rare cases of three-component systems in which the AFS-sets consist of more than 3 isolated subsets can also be treated by the inverse polygon inflation method. Also a "standard" three-subset AFS can be approximated by inverse polygon inflation. However, the standard polygon inflation method is faster for this case. Punctiform or line-shaped AFS-subsets cannot be detected. The progress of the approximation process of $\mathcal{M}_C$ by inverse polygon inflation for data set 2 is presented in figure 21. Figure 22 shows a flow chart of the inverse polygon inflation algorithm.

4.7. Ray casting for $s \geq 2$

Ray casting is a simple, general algorithm for AFS computations for any dimension $s \geq 2$ [35]. The method is stable for noisy data and can be used to calculate dimension-degenerated AFS subsets as line-intervals for $s = 3$, faces for $s \geq 4$ and so on. Ray casting is computationally faster than grid-search and of comparable computational costs as triangle enclosure. It is much slower than the geometric construction approaches or the polygon inflation method.

Ray casting works like radar scanning. A set of N rays is sent to uniformly distributed directions in the $(s-1)$-dimensional sphere. Polar coordinates (in 2D) or spherical coordinates (for 3 or more dimensions) can be used to construct the uniformly distributed directions. If v_i, $i = 1, \ldots, N$, are the direction vectors starting at the origin, then the points of intersections $R_i v_i$ with the boundary of $\mathcal{F}_S$ are computed. The R_i are real numbers. If $R_i v_i$ is a feasible point, then the smallest possible positive r_i is determined so that $r_i v_i$ is also a feasible point. Then the line segment $[r_i, R_i]v_i$ (where $[r_i, R_i]$ is the interval of real numbers from r_i to R_i) is the intersection of the AFS with the semi-infinite ray with the direction v_i. The theoretical basis for showing that this intersection has the form of a contiguous interval

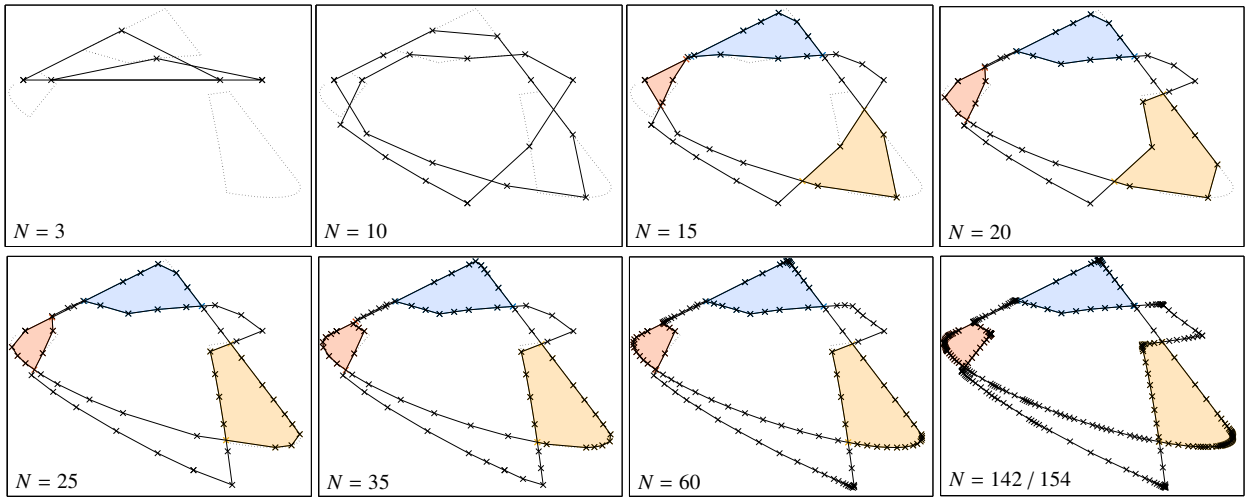

Figure 21: Progress of the inverse polygon inflation algorithm in order to compute $\mathcal{M}_C$ for data set 2. The set $\mathcal{F}_S$ and the subset $\mathcal{M}_S^*$ by (17) are computed by the polygon inflation procedure (black lines). Their intersection results in $\mathcal{M}_S$. The intermediate AFS approximation are colored. Nonrelevant parts of $\mathcal{M}_S^*$ are cut away. The number of current vertices is N.

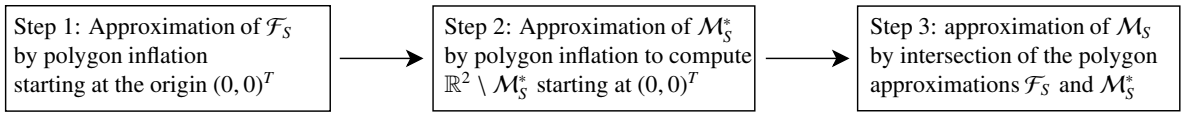

Figure 22: Flow chart of the inverse polygon inflation algorithm.

is theorem 3.4 (items 2 and 3). The resulting interval-shaped intersections of the AFS with the rays are clustered. The endpoints of neighboring intervals are interpolated within each cluster. This forms the boundary curve (or surface) of a subset of the AFS. The ray casting method is similar to the grid-search method but works with one less dimension. Figure 23 demonstrates an application of ray casting for the approximation of the spectral AFS for the data set 2 with a number of $N = 100$ rays.

Ray casting can be implemented with an adaptive control. Then areas are detected in which the ray density should be increased in order to avoid vertices not being detected. The adaptive strategy requires an initial ray casting approximation with N_{initial} rays. A vertex detection algorithm tries to identify intersecting edges and locally increases the resolution. The further refinement is stopped if only small local errors are detected. The adaptive strategy is very effective compared to a global, uniform refinement. Figure 24 illustrates the local refinement in the neighborhood of crucial vertices.

4.8. On the influence of the control parameters bounding small negative entries

As already pointed out in section 3.7, it appears to be a necessity to accept and to control small negative entries in the pure component factors C and S. This is especially the case for noisy data or in the presence of background signals or nonlinear effects violating the ideal bilinearity of the Lambert-Beer law. Otherwise, AFS approximation methods would be restricted to ideal model data. All generically numerical AFS approximation methods and also the dual Borgen plot construction can deal with noise and small negative matrix entries. With certain limitations this is also valid for the generalized Borgen plots approach. The size of the control parameters allow us to distinguish between acceptable deviations from the constraints $C, S \geq 0$ as well as $D = CS^T$ and unacceptably large deviations.

Even the classical ssq-function approach from [34, 49, 8, 50] can deal with noise and small negative matrix entries. The technique is to set the negative elements to zero and then to evaluate the reconstruction functional $\|D - CS^T\|$ that should be as small as possible. We use a different approach in this chapter, in [9, 10, 67, 68, 35, 58] and in the FACPACK software package, namely the parameters ε_C and ε_S control the relative size of negative entries as specified in equation (13).

A proper choice of these parameters is important in order to compute meaningful results. Generally, for increasing values ε_C and ε_S the area of the AFS increases. If these control parameters and thus the AFS are too small, then the chemically correct solution might not be contained in the AFS. Otherwise, if the control parameters together with

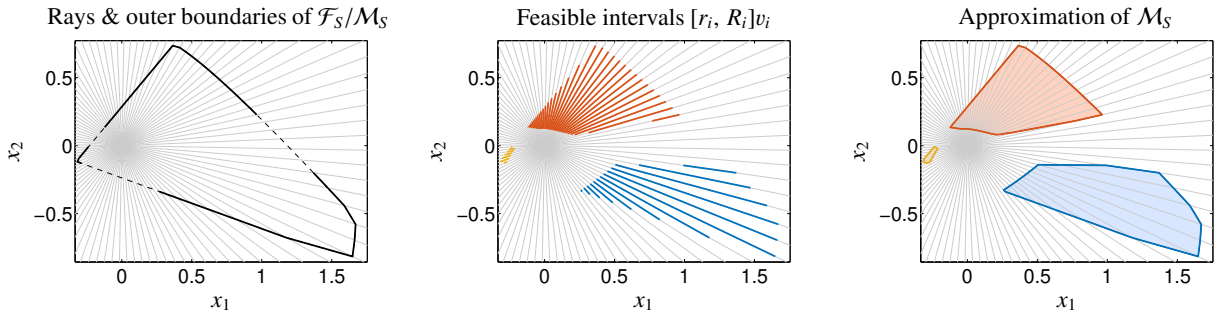

Figure 23: Application of the ray casting algorithm with $N = 100$ rays in order to approximate the spectral AFS for data set 2. Left: The equiangular rays starting at the origin are drawn in gray. The intersection of $\mathcal{M}_S$ with the boundary of $\mathcal{F}_S$ is marked by solid broken lines. The remaining parts of the boundary of $\mathcal{F}_S$ are drawn by black broken lines. Center: The intersection $[r_i, R_i]v_i$ of the ray with the direction v_i with the AFS $\mathcal{M}_S$ is marked in red, blue or ochre. Right: Piecewise linear interpolation of the end points of neighboring lines yields the boundary curve of an AFS subset.

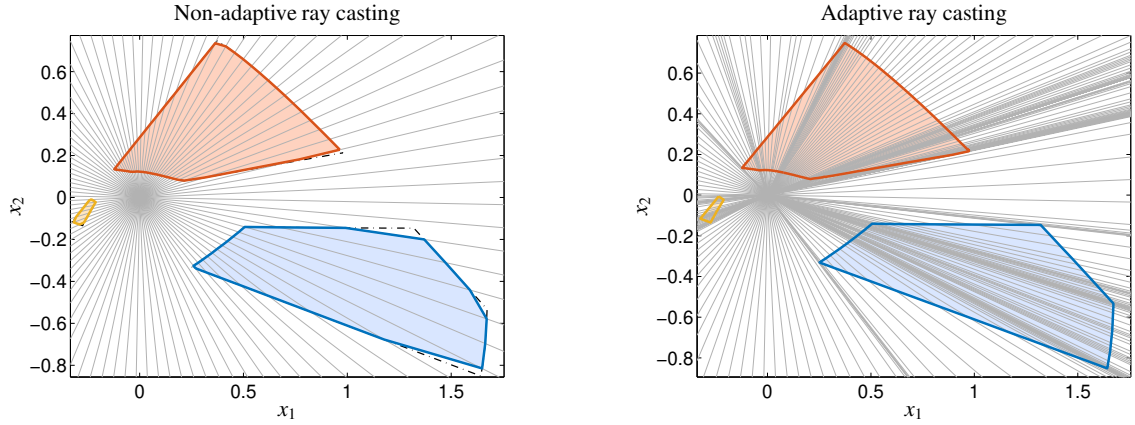

Figure 24: Ray casting with and without an adaptive refinement strategy to approximate the spectral AFS for data set 2. Left: Non-adaptive ray casting with a relatively low resolution by $N = 100$ equiangular rays tends to cut certain vertices. The dash-dotted black lines stand for the correct AFS. Especially the blue AFS subset shows truncated vertices. Right: Adaptive ray casting results in an improved approximation. The initial number of rays is $N_{\text{initial}} = 100$. The final approximation with a locally increased resolution and $N = 325$ rays.

the AFS are too large, then the information content of the AFS is small. Figure 25 shows the AFS-sets for various parameter choices for data set 3. Additionally, the true solutions (computed by using a kinetic hard model) are plotted. The first choice $\varepsilon_C = 3.88 \cdot 10^{-4}$ and $\varepsilon_S = 2.65 \cdot 10^{-3}$ is too restrictive as not all true factors are covered by the AFS. A proper selection for this data set appears to be $\varepsilon_C = \varepsilon_S = 0.006$.

4.9. Special AFS shapes

Next we determine the AFS for three nonnegative matrices with either the rank 3 or the rank 4. These matrices do not originate from a chemical reaction system but illustrate more general shapes of an AFS.

Example 4.1. *For the following three simple model problems the AFS $\mathcal{M}_C$ equals $\mathcal{M}_S$ apart from an axis scaling (that is the y_1-axis gets the scaling factor σ_1/σ_2 and y_2-axis is to be multiplied by σ_1/σ_3). For the first problem this fact is not surprising since D is a symmetric matrix. For the other two problems the structure of D^T is closely related to the structure of D.*

1. *This matrix is inspired by [69]. With*

$$C^T = S^T = \begin{pmatrix} \alpha & 1 & 1 & \alpha & \beta & \beta \\ 1 & \alpha & \beta & \beta & \alpha & 1 \\ \beta & \beta & \alpha & 1 & 1 & \alpha \end{pmatrix} \tag{18}$$

we get $D = CS^T$. We choose $\alpha = 0.5$ and $\beta = 0.075$. The AFS consists of six separated subsets, see figure 26. As $\sigma_2 = \sigma_3$, the orientation of the AFS depends on the choice of the singular vectors. The geometric Borgen plot construction, the grid-search method, the inverse polygon inflation algorithm and the ray casting approach

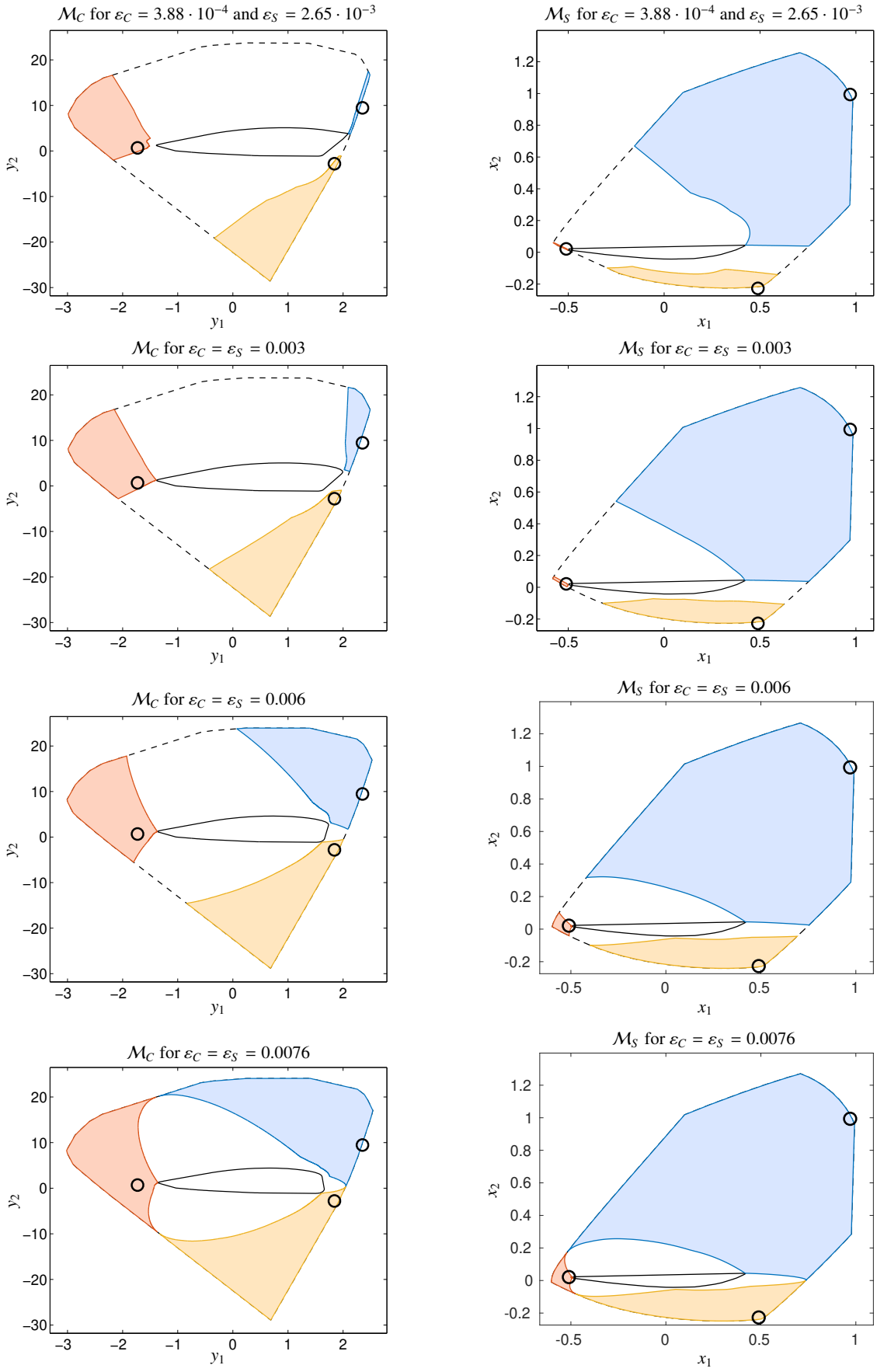

Figure 25: AFS computations for data set 3 and for different settings of the control parameters ε_C and ε_S, see equation (13). The outer polygons are drawn as black broken lines, the inner polygons by black solid lines and the low-dimensional representatives of the true solution (as computed by kinetic hard modeling) are marked by black circles. The AFS-sets grow for increasing values of ε_C, ε_S. If the control parameters are too small, then the true solutions are not necessarily contained in $\mathcal{M}_C$, respectively in $\mathcal{M}_S$.

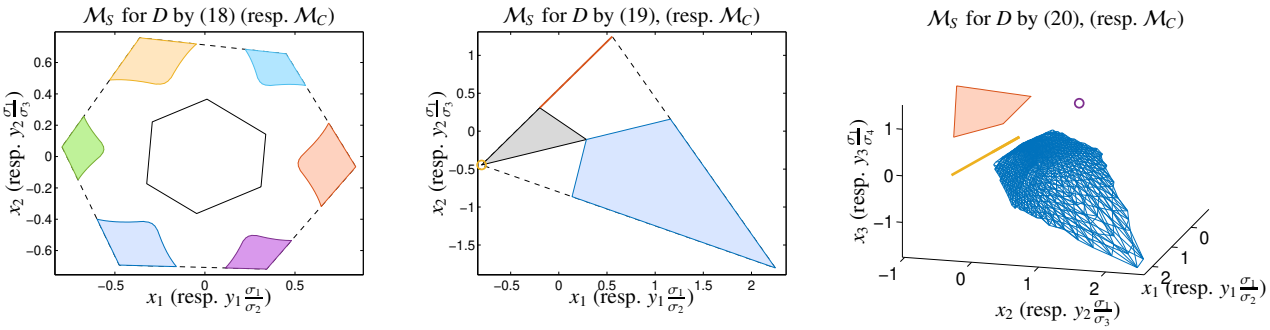

Figure 26: AFS shapes for the three matrices (18)–(20). For each of these matrices the AFS $\mathcal{M}_C$ equals $\mathcal{M}_S$ aside from scaling. The outer polygons are plotted by black broken lines and the inner polygons are drawn by black solid lines respectively by a gray area. Left: The AFS $\mathcal{M}_S$ for D by (18) with $\alpha = 0.5$ and $\beta = 0.075$. Due to $\sigma_2 = \sigma_3$ the orientation of the axes remains undetermined. Center: The AFS $\mathcal{M}_S$ for D by (19) with a 0-, 1- and 2-dimensional subset. Right: The AFS $\mathcal{M}_S$ for D by (20) consists of subsets with the dimensions 0 up to 3.

can be applied in order to compute this AFS. A successful application of triangle enclosure and direct polygon inflation would require algorithmic modifications.

2. *The nonnegative* 3×3 *matrix*

$$D = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \tag{19}$$

does not contain chemistry-related data. However, the AFS construction is possible and results in three isolated subsets. One subset is punctiform, one subset is line-shaped and the third subset is a two-dimensional area subset. The AFS construction succeeds with one of the geometric approaches (Borgen plots), the triangle enclosure method, the direct polygon inflation algorithm or the ray casting method. The methods of grid-search and inverse polygon inflation fail due to the dimension-degeneracy.

3. *The nonnegative* 4×4 *matrix*

$$D = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{20}$$

extends the construction principle underlying (19)*, namely of an all-1 upper triangular matrix, by increasing the dimension by* 1*. The AFS consists of a punctiform, a line-shaped, an area and a volume subset. The ray casting method or the sliced triangle enclosure method can be used in order to compute this AFS.*

5. AFS-computations with FACPACK

FACPACK is a software package written in MatLab (with a core of fast C- and FORTRAN-subroutines) for the computation of AFS-sets for two-, three- and four-component systems [11]. The software comes with a user-friendly graphical user interface (GUI) and is freely available on the FACPACK-homepage

http://www.math.uni-rostock.de/facpack/

Executable program codes of FACPACK are provided for Windows, Unix and macOS. The test data and the MatLab program code elements as presented in section 5.2 are also available on the FACPACK homepage.

5.1. Modules

The FACPACK software has a module structure that is explained in the following. The main modules are those on `AFS computation (2 and 3 components)`, `Duality/Complementarity & AFS (3 components)`, `Generalized Borgen plots` and `AFS computation (4 components)`. For all modules the user data has to be loaded from

*.mat-files. Therein the spectral mixture data has to be saved row-wise in the matrix D. Optionally, one can provide two grid vectors x and t. The vector x contains the coordinates on the abscissa of the spectra (typically frequency values or wavenumbers) and t contains the time values of the measurements. If x and t are not provided, then $x = (1, \ldots, n)$ and $t = (1, \ldots, k)$ is the default setting. For further explanations see [70, 11].

The polygon inflation algorithm and the inverse polygon inflation algorithm are implemented in the module `AFS-computation (2 and 3 components)`. Figure 27 serves as a quick reference for this module in order to compute the spectral AFS in application to data set 2. A special feature of these computations is the *live-view-mode*. The user can move the mouse pointer through the AFS and simultaneously the reconstructed (transformed) profiles are drawn. By pressing the left mouse button the profile under the current mouse pointer is locked. If one feasible solution is locked, then the remaining and reduced AFS-subsets can be computed by means of equality constraints [71, 72]. This process can be started by clicking the no. 1-button. Next a second profile can be locked in a similar way. Again, the reduced AFS can be computed with the no. 2-button.

The module `Duality/Complementarity & AFS (3 components)` is designed in order to compute the two dual AFS-sets $\mathcal{M}_S$ and $\mathcal{M}_C$ and to construct a complete factorization step-by-step in an interactive way. Figure 28 contains a quick reference for this module for data set 2. A triangle in the AFS plane containing $\mathcal{M}_S$ defines a complete factor S and therefore simultaneously determines the factor C. This triangle is directly related to a second triangle in the AFS plane that contains $\mathcal{M}_C$. The vertices of this triangle determine the columns of C. In this module the user can select and modify the vertices of the two triangles by selecting and moving the vertices by the computer mouse in the form of a live-view-mode. The current factors C and S are plotted simultaneously.

The geometric construction of the AFS is implemented in the module `Generalized Borgen plots`. In figure 29 a quick guide for this module is presented in application to the spectral AFS for data set 2. The module is very similar to the module `AFS-computation (2 and 3 components)`. Differences exist in regard to the axis scaling. One can either use a row-sum scaling (RSV-scaling) according to [6, 43] or the first-singular-vector scaling (FSV-scaling) as used in equation (7). Further, the control parameters on the extent of acceptable perturbations have to be selected in an absolute way.

Finally, the ray casting algorithm for four-component systems is implemented in the module `AFS-computation (4 components)`. Figure 30 illustrates a quick guide for this module in order to compute the spectral AFS for the example four_components/example5.mat that is supplied by the FACPACK software. The *live-view-mode* works if a 3D-mouse is used and the program runs under Windows. Sliders are available for the 3 coordinates. The isolated subsets of the AFS can be plotted separately.

5.2. Useful hints how to get access to the FACPACK results and figures

FACPACK provides the user with the possibility of saving the computational results in individual files. These files can either be loaded in the command window or can be used for an independent post-processing of the results. The variables are saved in the form of structure-formatted data.

If for example the `AFS-computation (2 and 3 components)`-module has been used for a three-component system and the results have been saved by the `Save all`-button, then the AFS is saved as a cell array. The boundary curves of the AFS-subsets are stored in AFS{1}, AFS{2} and AFS{3}. Then AFS{i}(:,1) refers to the x_1-coordinates and AFS{i}(:,2) to the x_2-coordinates. The sets of feasible profiles are saved in the cell array TransformedAFS and can be accessed in an analogous way.

Next we present some useful MatLab-commands for the presentation and modification of the FACPACK results.

Plot an AFS that consists of three separated subsets:

```
COL = [.85 .85 1; .85 1 .85; 1 .85 .85];
col = [0 0 1; 0 .5 0; 1 0 0];
load results_facpack_afs_USERDATA.mat
for i=1:3
    fill(AFS{i}(:,1), AFS{i}(:,2), COL(i,:), 'linewidth', 1, 'edgecolor', col(i,:));
    hold on;
end
hold off
axis tight
xl = get(gca, 'xlim');
yl = get(gca, 'ylim');
set(gca, 'fontsize', 22, 'xlim', 1.05*xl, 'ylim', 1.05*yl)
```

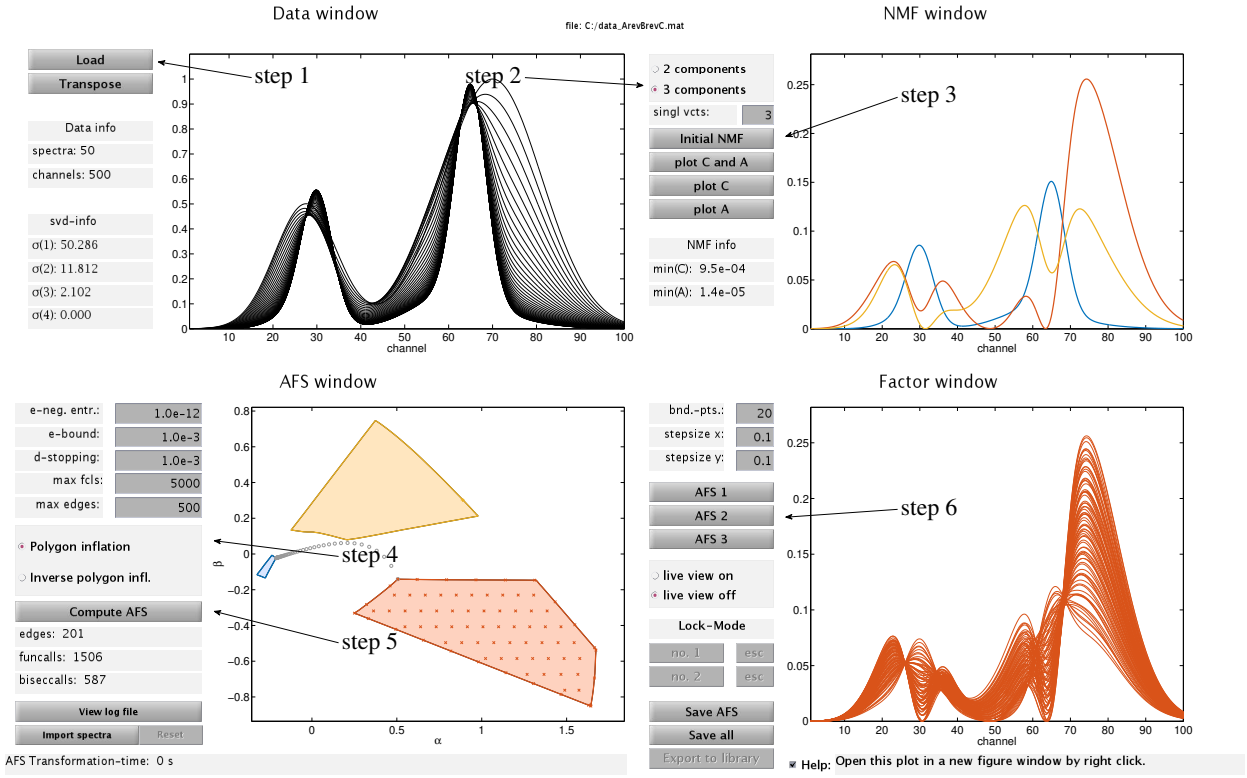

Figure 27: Quick reference in six steps for the module `AFS-computation (2 and 3 components)` that is an implementation of the polygon inflation algorithm. Data loading (step 1). Select the number of components, here $s = 3$ (step 2) and compute an initial NMF (step 3). Select the polygon inflation method (step 4) and compute the spectral AFS (step 5). Show the corresponding feasible profiles for one of the AFS subsets (step 6). The AFS for the concentration factor can be computed by pushing the Transpose button after data loading.

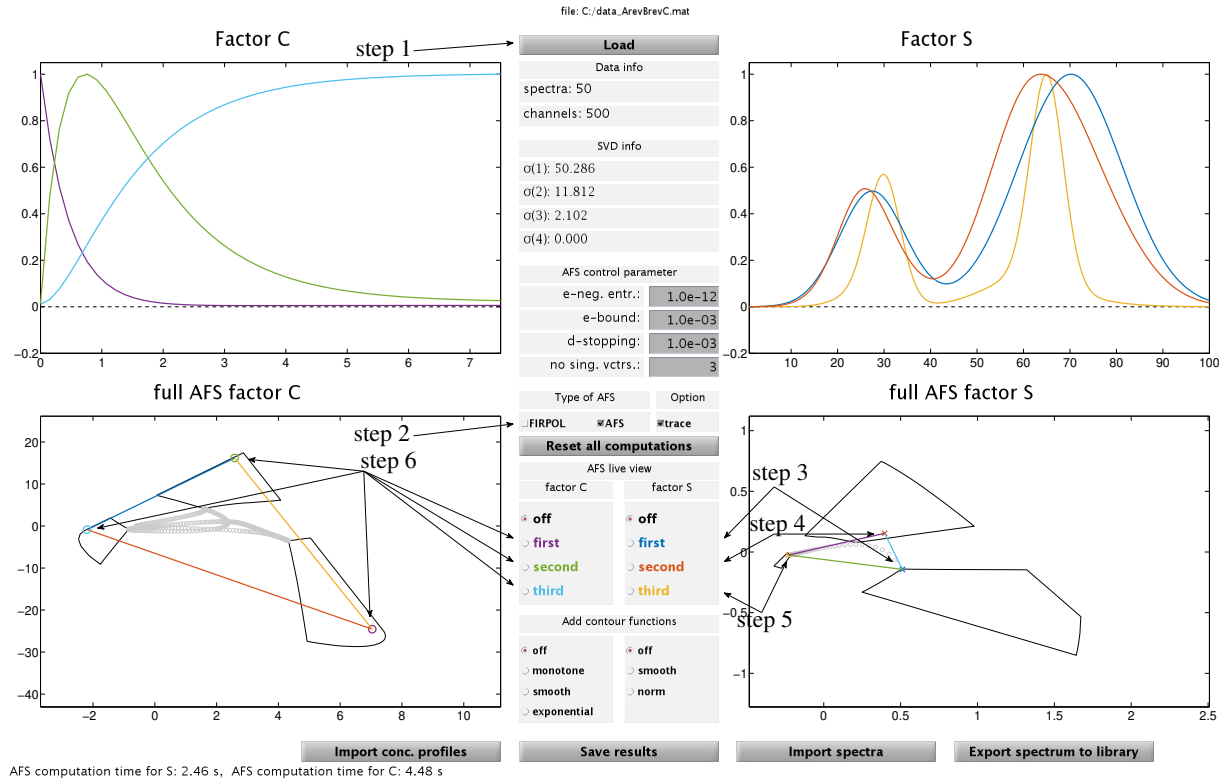

Figure 28: Quick reference in six steps for the module `Duality/Complementarity & AFS (3 components)`. Load the data (step 1). Compute the two dual AFS-sets (step 2). Select a first profile by moving the mouse pointer through the spectral AFS and lock a proper spectral profile (step 3). Select a second (step 4) and a third (step 5) profile in an analogous way. Any concentration profile or spectral profile can be changed in order to modify the pure component factorization (step 6).

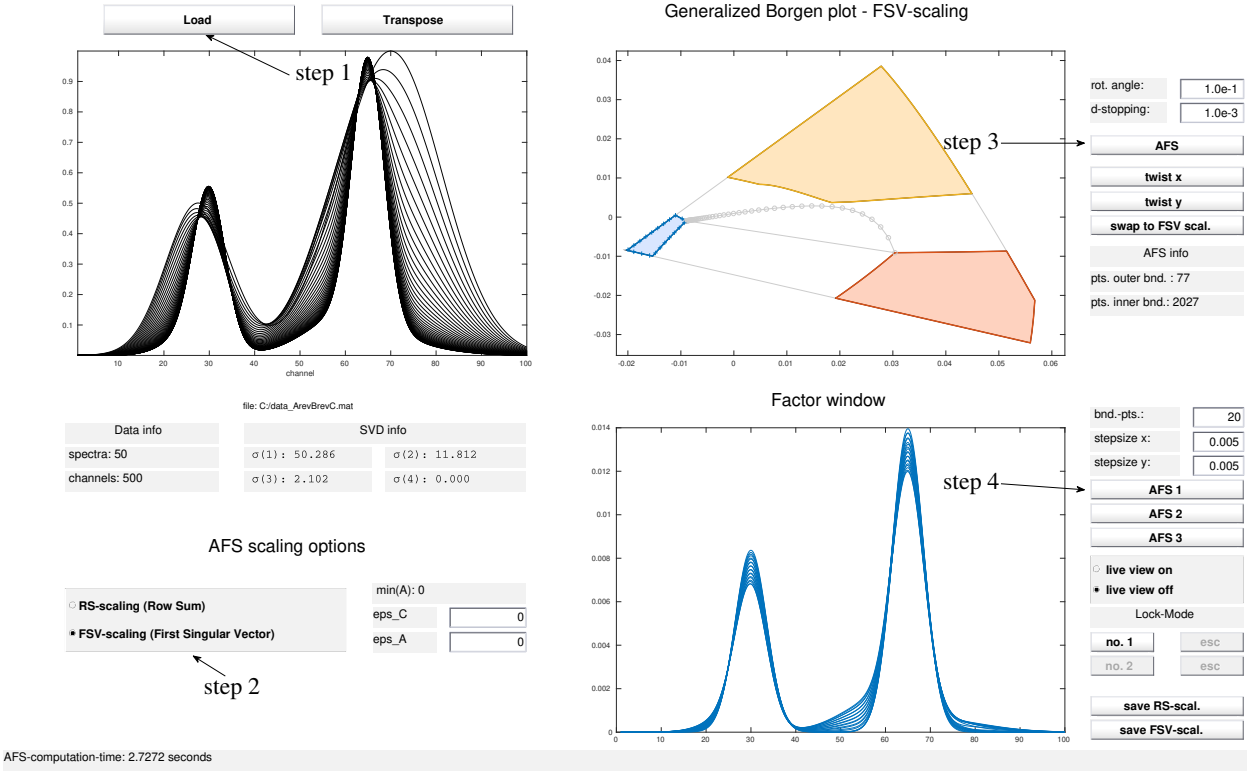

Figure 29: Quick reference in four steps for the module `Generalized Borgen plots` that implements the geometric AFS construction. Load the data (step 1). Select the AFS scaling (step 2) and compute the spectral AFS (step 3). Plot the corresponding feasible profiles for one of the subsets of the AFS (step 4).

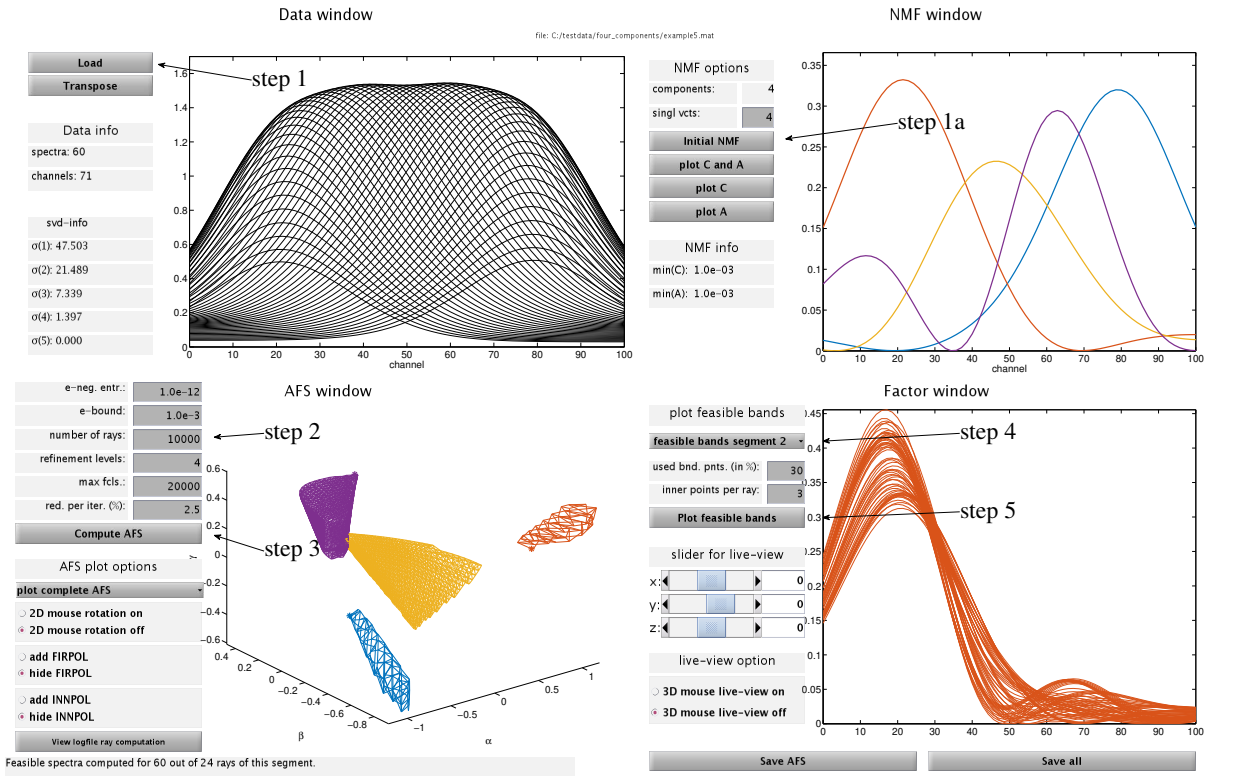

Figure 30: Quick reference in five steps for the module `AFS computation (4 components)` with the ray casting algorithm. Load the data (step 1). An initial NMF is computed automatically (step 1a). Select a proper number of rays (step 2) and compute the spectral AFS (step 4). Select a subset (step 4) in order to plot the corresponding feasible profiles (step 5). This demonstration uses example four_components/example5.mat that is part of the FACPACK software.

```
set(gcf, 'paperpositionmode', 'auto')
print('-depsc', '-loose', 'USERPICTURE')
```

Plot an AFS that is a topologically connected set containing a hole around the origin:

```
load results_facpack_afs_USERDATA.mat
fill(AFS{1}(:,1), AFS{1}(:,2), [.7 .7 .7], 'linewidth', 1, 'edgecolor', 'k')
hold on;
fill(AFS{2}(:,1), AFS{2}(:,2), [1 1 1], 'linewidth', 1, 'edgecolor', 'k')
hold off
axis tight
xl = get(gca, 'xlim');
yl = get(gca, 'ylim');
set(gca, 'fontsize', 22, 'xlim', 1.05*xl, 'ylim', 1.05*yl)
set(gcf, 'paperpositionmode', 'auto')
print('-depsc', '-loose', 'USERPICTURE')
```

Plot the AFS and mark imported spectral profiles in the AFS:

```
COL = [.85 .85 1; .85 1 .85; 1 .85 .85];
col = [0 0 1; 0 .5 0; 1 0 0];
load results_facpack_afs_USERDATA.mat
load EXTERNALFILE % this file contains in the variable S columnwise spectral profiles
tt = S'*V(:,1:3);
tt = diag(1./tt(:,1))*tt;
for i=1:3
    fill(AFS{i}(:,1), AFS{i}(:,2), COL(i,:), 'linewidth', 1, 'edgecolor', col(i,:));
    hold on;
end
plot(tt(:,2), tt(:,3), 'kx', 'markersize', 10, 'linewidth', 1)
hold off
axis tight
xl = get(gca, 'xlim');
yl = get(gca, 'ylim');
set(gca, 'fontsize', 22, 'xlim', 1.05*xl, 'ylim', 1.05*yl)
set(gcf, 'paperpositionmode', 'auto')
print('-depsc', '-loose', 'USERPICTURE')
```

Plot the feasible profiles (for an AFS consisting of three isolated subsets):

```
col = [0 0 1;0 .5 0;1 0 0;0 1 1];
load results_facpack_all_USERDATA.mat
for i=1:3
    figure(i)
    plot(x, TransformedAFS{i}, 'color',col(i,:))
    axis tight
    yl = get(gca, 'ylim');
    set(gca, 'fontsize', 22, 'xlim', x([1 end]), 'ylim', [min(0, 1.05*yl(1)) 1.05*yl(2)])
    set(gcf,'paperpositionmode','auto')
    print('-depsc','-loose',['USERPIC_FP_',num2str(i)]);
end
```

Computation and plot of the inner and outer polygons:

```
load USERDATA.mat
% the spectral data matrix is D with the dimensions (k x n)
% k is the number of spectra, n is the number of channels
% s is the number of components
s = 3;
[U,S,V] = svds(D,s);
for i=1:s
    if min(V(:,i)) < -max(V(:,i))
        V(:,i) = -V(:,i);
```

```
        U(:,i) = -U(:,i);
    end
end
a = diag(1./U(:,1)/S(1,1))*U(:,2:3)*S(2:3,2:3);
b = diag(1./V(:,1))*V(:,2:3);
% computation of the indices of a resp. b that represent vertices of the inner polygons
cips = convhull(a(:,1),a(:,2));
cipc = convhull(b(:,1),b(:,2));
% computation of the inner polygons
IPS = a(cips,:);
IPC = b(cipc,:);
% computation of outer polygon (FIRPOL) for S by using duality and IPC
FPS = [];
m = size(IPC,1);
for i=1:m-1
    X = [IPC(i,:); IPC(i+1,:)];
    FPS(end+1,:) = X\[-1;-1];
end
FPS(end+1,:) = FPS(1,:);
% computation of outer polygon (FIRPOL) for C by using duality and IPS
FPC = [];
m = size(IPS,1);
for i=1:m-1
    X = [IPS(i,:); IPS(i+1,:)];
    FPC(end+1,:) = X\[-1;-1];
end
FPC(end+1,:) = FPC(1,:);
% plot outer polygon and inner polygon for S
figure(1)
plot(FPS(:,1), FPS(:,2), IPS(:,1), IPS(:,2), 'linewidth', 1)
% plot outer polygon and inner polygon for C
figure(2)
plot(FPC(:,1), FPC(:,2), IPC(:,1), IPC(:,2), 'linewidth', 1)
```

6. Numerical results and comparative analysis

In a comparative analysis the feasible concentration profiles and the feasible spectral profiles are analyzed for the data sets as introduced in section 1.1. The computation of the various areas of feasible solutions is the first, intermediate step. The associated feasible profiles are determined for each subset of the AFS. To this end, each AFS subset is covered by a regular grid. Further, the boundary curve is covered by additional equidistant grid points. The total number of nodes of the grid plus the number of nodes on the boundary curve is m. These nodes $z_1, \ldots, z_m \in \mathbb{R}^{s-1}$ represent the spectral profiles according to

$$a^{(i)} = (1, z_i^T)V^T, \qquad i = 1, \ldots, m.$$

First, figure 31 shows the AFS and the bands of feasible solutions for the two-component data set 1. Figure 32 presents the analogous results for data set 2. In each case, the AFS subsets show the grid that is used for the band evaluation. Finally, for data set 3 the feasible profiles are computed for the AFS-sets as plotted in the third row of figure 25. The control parameters of this calculation are $\varepsilon_C = \varepsilon_S = 6 \cdot 10^{-3}$. The bands of feasible profiles are presented in figure 33.

6.1. MCR-BANDS results

The MCR-BANDS software supplies extremal concentration profiles and spectra by minimizing or maximizing the signal contribution function (5), see [31, 32, 21]. For an s-component system $2s$ spectral profiles and $2s$ concentration profiles are computed. Here we apply MCR-BANDS to data set 2. For each of the resulting profiles we compute the low-dimensional AFS representatives. Figure 32 shows in its two lower rows of plots the bands of solutions for equally distributed points in the AFS together with the extremal profiles by MCR-BANDS. The equally distributed points in the AFS are marked by crosses in the upper row of plots. The MCR-BANDS profiles are plotted as black broken lines. These profiles represent useful approximate enclosures for the bands of solutions. However, in some

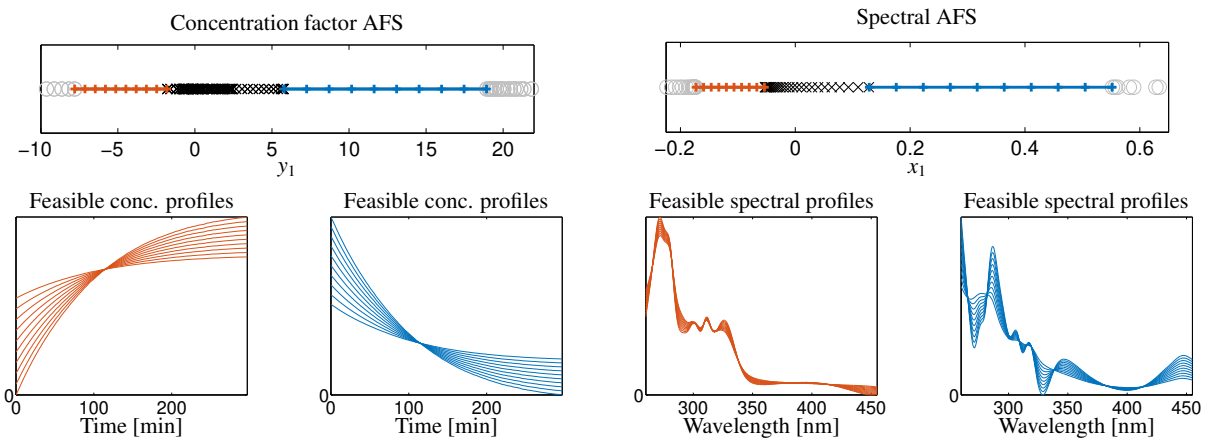

Figure 31: Analysis of the feasible profiles for the two-component data set 1. The upper row shows the AFS $\mathcal{M}_C$ for the concentration factor and the corresponding AFS for the spectral factor $\mathcal{M}_S$. The gray circles serve to define the outer boundary and the black crosses are the data-representing vectors. The red and blue line markers in the AFS mark the positions for which the spectral profiles and concentration profiles are evaluated. These profiles are drawn in the lower row of figures.

regions, e.g. the upper bound of the spectral profiles for Y in the $\lambda \approx 80...100$ interval, the approximation is poor. Additionally, figure 34 shows where the representatives of the MCR-BANDS profiles are located in the AFS sets. In general, the MCR-BANDS profiles have the tendency to be represented by points that are located on the boundary of the AFS. However, the MCR-BANDS maximum profile for the spectrum of Y is associated with a point close to the boundary, see also [73, 55]. In figure 34 faint markers illustrate the convergence history of the iterative optimization in MCR-BANDS.

6.2. Comparison of computational results

Finally, we compare some of the various methods with respect to their precision and the computational costs. This comparison is restricted to the three-component model data set 2. The reference solution is constructed geometrically (Borgen plots) with a number of $N = 12\,610$ tangents.

For a fair comparison each of the methods is implemented in fast C-code. Furthermore, we apply the same program routine for the feasibility check of points in the AFS plane. For this purpose we consistently use the objective function $F(x)$ by (16) and not the ssq-function. All computations have been performed on a standard PC with an Intel i7-6700 CPU with 3.40GHz and 16 GB RAM. Only one processor core has been used.

The tables 1-6 contain the numerical data for the grid-search method, the Borgen plot approach, the triangle enclosure method, the polygon inflation algorithm, the inverse polygon inflation algorithm and the ray casting approach with adaptive control. We compare the accuracy of the boundary approximation for each of the three AFS subsets (subset 1 belongs to component X, subset 2 to Y and subset 3 to Z). Furthermore, figure 35 compares the methods in a log-log graph with respect to the boundary approximation accuracy and the computation time. The result of these comparisons for the model data set 2 is as follows: For this noise-free data, the geometric Borgen plot construction is the fastest and most accurate method. Polygon inflation follows on the next position. Grid search and triangle enclosure require more computation time in order to achieve a comparable accuracy. The adaptive ray casting method has a mid-position.

7. Summary

The reliability of multivariate curve resolution methods in chemometrics is largely determined by an appropriate approach to deal with the ambiguity underlying pure component factorizations. The systematic analysis of this ambiguity leads to the area of feasible solutions as a low-dimensional representation of the possible factors only with respect to the nonnegativity constraint.

The wide range of possible methods to construct or to compute these ranges of feasible solutions is reviewed in this chapter. Many of the fundamental properties of the AFS have been highlighted. The various algorithms and their program implementations for the computation of feasible bands and/or the AFS enable an unbiased view on the annoying ambiguity problem.

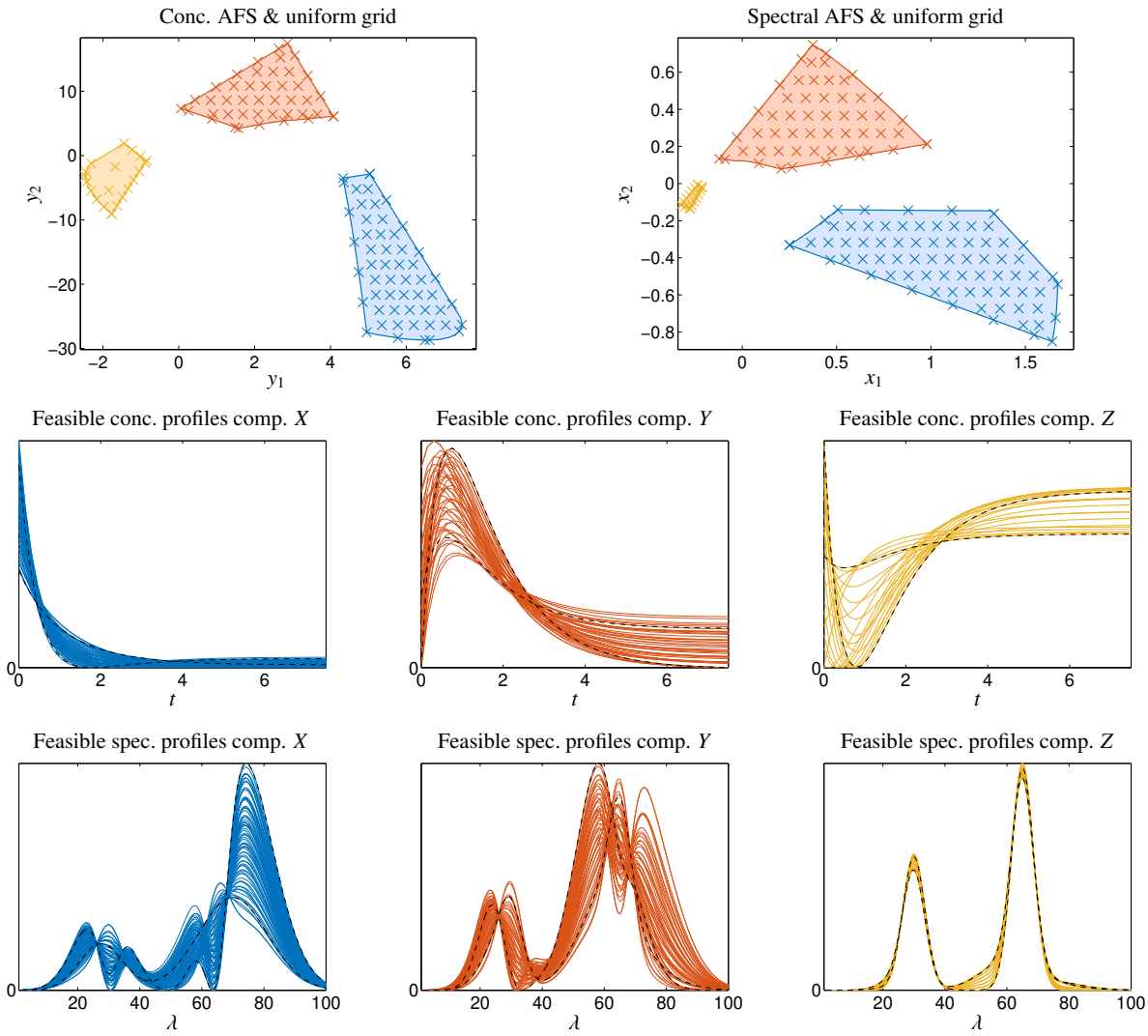

Figure 32: For data set 2 this figure shows the AFS-sets $\mathcal{M}_C$ and $\mathcal{M}_S$ in the top row. Colored crosses mark the AFS elements for which the associated concentration profiles are drawn in the middle row of plots. The associated spectral profiles are drawn in the lower row. Additionally, the MCR-BANDS profiles are plotted as black broken lines.

We hope that the FACPACK software can substantially support the user in his tour through the chemical data analysis. We explicitly recommend the application of the program segments, as introduced in section 5.2, for the data post-processing. These programs allow the user a deep access to all computational results of FACPACK for his personal research purposes and for graphical visualization tasks.

References

[1] E. C. Cherry. Some experiments on the recognition of speech, with one and with two ears. *J. Acoust. Soc. Am.*, 25(5):975–979, 1953.

[2] R. Kellner, J.-M. Mermet, M. Otto, M. Valcárcel, and H. M. Widmer, editors. *Analytical chemistry*. Wiley-VCH, Weinheim, 2004.

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

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

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

[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] H. Abdollahi and R. Tauler. Uniqueness and rotation ambiguities in multivariate curve resolution methods. *Chemom. Intell. Lab. Syst.*, 108(2):100–111, 2011.

[8] A. Golshan, H. Abdollahi, and M. Maeder. Resolution of rotational ambiguity for three-component systems. *Anal. Chem.*, 83(3):836–841, 2011.

[9] 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.

[10] 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.

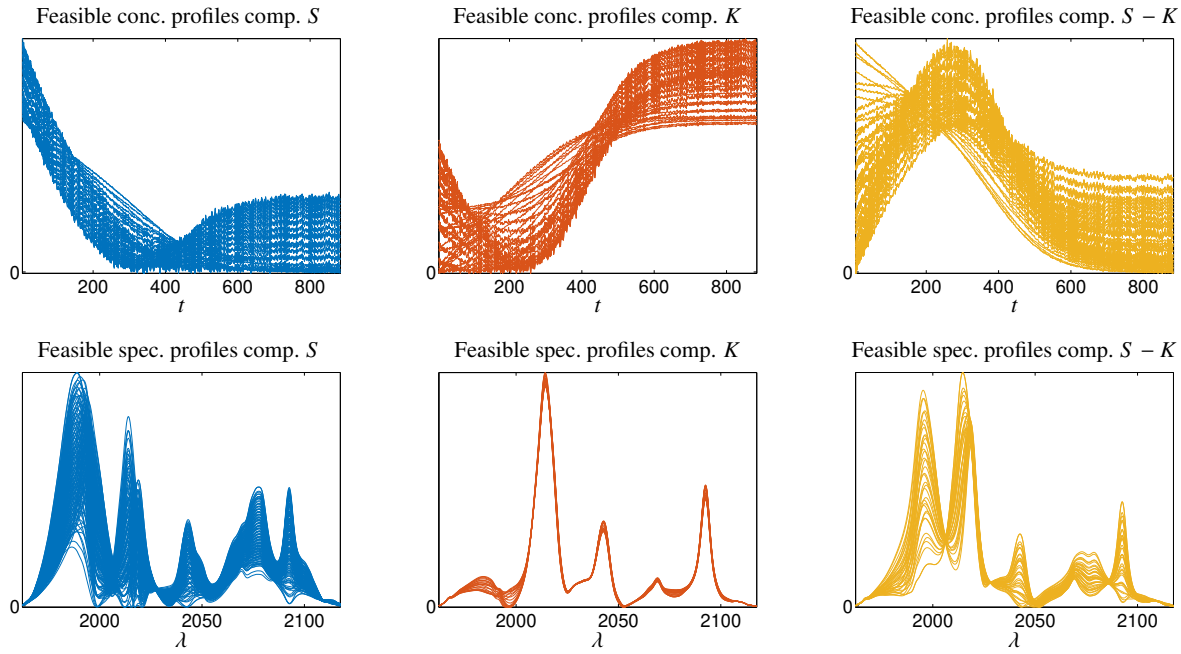

Figure 33: Feasible profiles for data set 3 with $\varepsilon_C = \varepsilon_S = 6 \cdot 10^{-3}$.

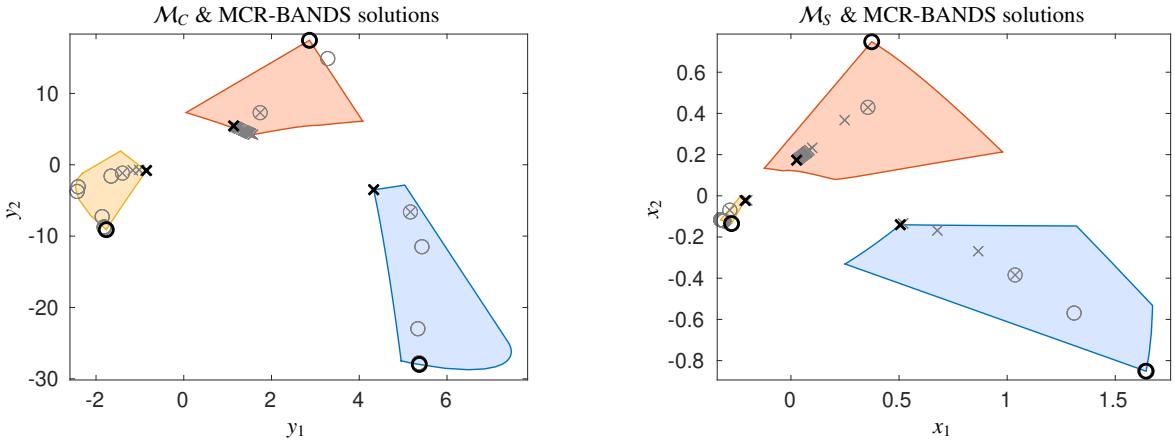

Figure 34: For data set 2 this plot shows the MCR-BANDS representatives in the AFS. The maxima of the signal contribution function are marked by × symbols. Small circles are used for the minima. The gray markers illustrate the convergence history of MCR-BANDS.

[11] M. Sawall, A. Moog, and K. Neymeyr. FACPACK: A software for the computation of multi-component factorizations and the area of feasible solutions, Revision 1.3. FACPACK homepage: http://www.math.uni-rostock.de/facpack/, 2018.

[12] C. Fischer, S. Schulz, H.-J. Drexler, C. Selle, M. Lotz, M. Sawall, K. Neymeyr, and D. Heller. The Influence of Substituents in Diphosphine Ligands on the Hydrogenation Activity and Selectivity of the Corresponding Rhodium Complexes as Exemplified by ButiPhane. *ChemCatChem*, 4(1):81–88, 2012.

[13] 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.

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

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

[16] 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.

[17] H. Schröder, M. Sawall, C. Kubis, D. Selent, D. Hess, R. Franke, A. Börner, and K. Neymeyr. On the ambiguity of the reaction rate constants in multivariate curve resolution for reversible first-order reaction systems. *Anal. Chim. Acta*, 927:21–34, 2016.

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

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

[20] V. P. Pauca, J. Piper, and R.J. Plemmons. Nonnegative matrix factorization for spectral data analysis. *Linear Algebra Appl.*, 416(1):29–47, 2006.

[21] 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.

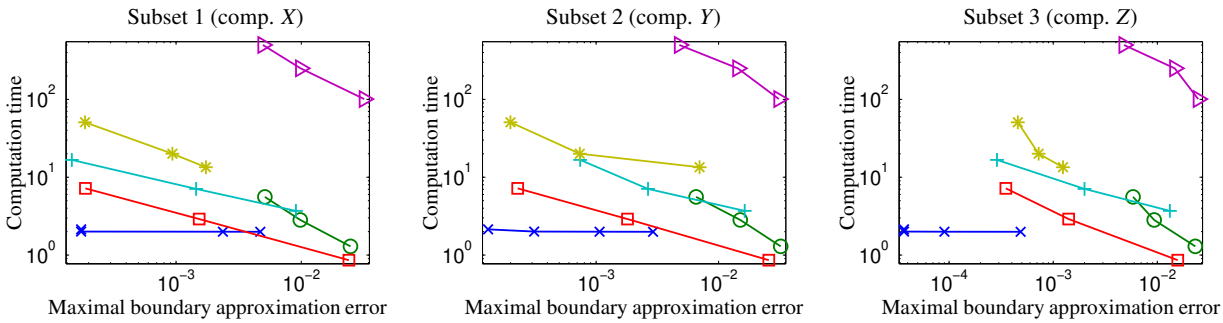

Figure 35: Comparison of the various AFS-computation methods for $\mathcal{M}_S$ and data set 2. The computation time is plotted as a function of the calculated boundary approximation error in a log-log plot. The symbol-method relation is as follows: geometric construction (×), grid-search (▷), triangle enclosure (○), polygon inflation (□), inverse polygon inflation (+), adaptive ray casting (✳). For this model data set the geometric AFS construction by Borgen plots is clearly the best method. For these plots the numerical values are taken from the tables 1-6.

[22] 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.

[23] J. Jaumot, A. de Juan, and R. Tauler. MCR-ALS GUI 2.0: new features and applications. *Chemom. Intell. Lab. Syst.*, 140:1–12, 2015.

[24] E. Widjaja, C. Li, and M. Garland. Semi-batch homogeneous catalytic in-situ spectroscopic aata. FTIR spectral reconstructions using Band-Target Entropy Minimization (BTEM) without spectral preconditioning. *Organometallics*, 21:1991–1997, 2002.

[25] 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.

[26] M. Sawall, C. Kubis, E. Barsch, D. Selent, A. Börner, and K. Neymeyr. Peak group analysis for the extraction of pure component spectra. *J. Iran. Chem. Soc.*, 13(2):191–205, 2016.

[27] D.D. Lee and H.S. Seung. Algorithms for non-negative matrix factorization. *Advances in neural information processing systems*, 13:556–562, 2001.

[28] H. Kim and H. Park. Algorithms for nonnegative natrix and tensor factorizations: A unified view based on block coordinate descent framework. *J. Global Optim.*, 58(2):285–319, 2014.

[29] H. Gampp, M. Maeder, C. J. Meyer, and A. D. Zuberbühler. Calculation of equilibrium constants from multiwavelength spectroscopic data–II: SPECFIT: two user-friendly programs in basic and standard FORTRAN 77. *Talanta*, 32(4):257–264, 1985.

[30] M. Maeder and P. King. Reactlab, 2009.

[31] 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.

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

[33] E. Kriesten, F. Alsmeyer, A. Bardow, and W. Marquardt. Fully automated indirect hard modeling of mixture spectra. *Chemom. Intell. Lab. Syst.*, 91:181–193, 2008.

[34] 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.

[35] M. Sawall and K. Neymeyr. A ray casting method for the computation of the area of feasible solutions for multicomponent systems: Theory, applications and FACPACK-implementation. *Anal. Chim. Acta*, 960:40–52, 2017.

[36] C. Eckard and G. Young. The approximation of one matrix by another of lower rank. *Psychometrika*, 1(3):211–218, 1936.

[37] G.W. Stewart. On the early history of the singular value decomposition. *SIAM Review*, 35:551–556, 1993.

[38] J. E. Jackson. *A user's guide to principal components*. Wiley, 1991.

[39] I. T. Jolliffe. *Principal component analysis, 2nd edition*. Springer, 2002.

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

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

[42] A. Berman and R. J. Plemmons. *Nonnegative matrices in the mathematical sciences*, volume 9 of *Classics Appl. Math.* SIAM, Philadelphia, 1994.

[43] 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.

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

[45] A. Jürß, M. Sawall, and K. Neymeyr. On generalized Borgen plots. II: The line-moving algorithm and its numerical implementation. *J. Chemom.*, 30:636–650, 2016.

[46] M. Sawall, A. Jürß, H. Schröder, and K. Neymeyr. Simultaneous construction of dual Borgen plots. I: The case of noise-free data. *J. Chemom.*, 31:e2954, 2017.

[47] K. Sasaki, S. Kawata, and S. Minami. Constrained nonlinear method for estimating component spectra from multicomponent mixtures. *Applied Optics*, 22(22):3599–3603, 1983.

[48] A. Meister. Estimation of component spectra by the principal components method. *Anal. Chim. Acta*, 161:149–161, 1984.

[49] 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.

[50] 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.

[51] A. Golshan, H. Abdollahi, S. Beyramysoltan, M. Maeder, K. Neymeyr, R. Rajkó, M. Sawall, and R. Tauler. A review of recent methods for the determination of ranges of feasible solutions resulting from soft modelling analyses of multivariate data. *Anal. Chim. Acta*, 911:1–13, 2016.

[52] M. Sawall, A. Jürß, H. Schröder, and K. Neymeyr. *On the analysis and computation of the area of feasible solutions for two-, three- and four-component systems*, volume 30 of Data Handling in Science and Technology, "Resolving Spectral Mixtures", Ed. C. Ruckebusch, chapter 5, pages 135–184. Elsevier, Cambridge, 2016.
[53] R.S. Varga. *Matrix Iterative Analysis*. Springer Series in Computational Mathematics. Springer Berlin Heidelberg, 1999.
[54] K. Neymeyr and M. Sawall. On the set of solutions of the nonnegative matrix factorization problem. *SIAM J. Matrix Anal. Appl.*, 39:1049–1069, 2018.
[55] 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.
[56] R.C. Henry. Duality in multivariate receptor models. *Chemom. Intell. Lab. Syst.*, 77(1-2):59–63, 2005.
[57] R. Rajkó. Natural duality in minimal constrained self modeling curve resolution. *J. Chemom.*, 20(3-4):164–169, 2006.
[58] M. Sawall, A. Moog, C. Kubis, H. Schröder, D. Selent, R. Franke, A. Brächer, A. Börner, and K. Neymeyr. Simultaneous construction of dual Borgen plots. II: Algorithmic enhancement for applications to noisy spectral data. *J. Chemom.*, 32:e3012, 2018.
[59] 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.
[60] 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.
[61] M. Sawall, C. Kubis, R. Franke, D. Hess, D. Selent, A. Börner, and K. Neymeyr. How to apply the complementarity and coupling theorems in MCR methods: Practical implementation and application to the Rhodium-catalyzed hydroformylation. *ACS Catal.*, 4:2836–2843, 2014.
[62] K. Neymeyr and M. Sawall. On an SVD-free approach to the complementarity and coupling theory: A note on the elimination of unknowns in sums of dyadic products. *J. Chemom.*, 30:30–36, 2016.
[63] J. Dennis, D. Gay, and R. Welsch. An adaptive nonlinear least-squares algorithm. *ACM Transactions on Mathematical Software*, 7:348–368, 1981.
[64] J. Dennis, D. Gay, and R. Welsch. Algorithm 573: An adaptive nonlinear least-squares algorithm. *ACM Transactions on Mathematical Software*, 7:369–383, 1981.
[65] A.N. Skvortsov. Estimation of rotation ambiguity in multivariate curve resolution with charged particle swarm optimization (cPSO-MCR). *J. Chemom.*, 28(10):727–739, 2014.
[66] S. Ghaheri, S. Masoum, and A. Gholami. Resolving of challenging gas chromatography-mass spectrometry peak clusters in fragrance samples using multicomponent factorization approaches based on polygon inflation algorithm. *J. Chromatogr. A*, 1429:317–328, 2016.
[67] M. Sawall, N. Rahimdoust, C. Kubis, H. Schröder, D. Selent, D. Hess, H. Abdollahi, R. Franke, Börner A., and K. Neymeyr. Soft constraints for reducing the intrinsic rotational ambiguity of the area of feasible solutions. *Chemom. Intell. Lab. Syst.*, 149, Part A:140–150, 2015.
[68] N. Rahimdoust, M. Sawall, K. Neymeyr, and H. Abdollahi. Investigating the effect of flexible constraints on the accuracy of self-modeling curve resolution methods in the presence of perturbations. *J. Chemom.*, 30(5):252–267, 2016.
[69] H. Laurberg, M.G. Christensen, M.D. Plumbley, L.K. Hansen, and S.H. Jensen. Theorems on Positive Data: On the Uniqueness of NMF. *Computational Intelligence and Neuroscience*, 2008:9 pages, 2008.
[70] 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.
[71] 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.
[72] 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.
[73] M. Sawall, A. Golshan, R. Tauler, and K. Neymeyr. Does the signal contribution function attain its extrema in boundary points of the Borgen plot? Technical report, 2019.

Rotation angle [°]	Number of tangents	Computation time [s]	Boundary accuracy subset 1	subset 2	subset 3	Number of vertices
1.00	459	1.99	$4.711 \cdot 10^{-3}$	$2.917 \cdot 10^{-3}$	$4.866 \cdot 10^{-4}$	99
0.50	819	2.00	$2.367 \cdot 10^{-3}$	$1.066 \cdot 10^{-3}$	$9.023 \cdot 10^{-5}$	162
0.10	3699	2.00	$1.746 \cdot 10^{-4}$	$3.118 \cdot 10^{-4}$	$3.689 \cdot 10^{-5}$	667
0.05	7299	2.14	$1.746 \cdot 10^{-4}$	$1.310 \cdot 10^{-4}$	$3.689 \cdot 10^{-5}$	129

Table 1: Geometric construction of $\mathcal{M}_S$ by triangle rotation according to the Borgen plot construction for data set 2. The rotation angle varies between 1 and 0.05 degrees.

Step size h_{x_1}, h_{x_2}	Grid	Computation time [s]	Boundary accuracy subset 1	subset 2	subset 3	Number of vertices
$2.5 \cdot 10^{-2}$	101×73	13.53	$3.183 \cdot 10^{-2}$	$3.107 \cdot 10^{-2}$	$2.469 \cdot 10^{-2}$	254
$1.0 \cdot 10^{-2}$	251×181	78.81	$9.879 \cdot 10^{-3}$	$1.433 \cdot 10^{-2}$	$1.469 \cdot 10^{-2}$	654
$5.0 \cdot 10^{-3}$	501×361	280.23	$4.941 \cdot 10^{-3}$	$4.785 \cdot 10^{-3}$	$4.785 \cdot 10^{-3}$	1315
$1.0 \cdot 10^{-3}$	2501×1801	6142.81	$9.930 \cdot 10^{-4}$	$9.643 \cdot 10^{-4}$	$9.045 \cdot 10^{-4}$	6623

Table 2: Application of the grid-search approach in order to approximate $\mathcal{M}_S$ for data set 2. The results are computed for several grid resolutions.

Edge length a	Computation time [s]	Boundary accuracy subset 1	subset 2	subset 3	Number of triangles
$2.5 \cdot 10^{-2}$	1.29	$2.474 \cdot 10^{-2}$	$3.217 \cdot 10^{-2}$	$2.317 \cdot 10^{-2}$	565
$1.0 \cdot 10^{-2}$	2.83	$9.830 \cdot 10^{-3}$	$1.504 \cdot 10^{-2}$	$9.329 \cdot 10^{-3}$	1429
$5.0 \cdot 10^{-3}$	5.58	$5.122 \cdot 10^{-3}$	$6.550 \cdot 10^{-3}$	$5.846 \cdot 10^{-3}$	2865
$1.0 \cdot 10^{-3}$	27.43	$1.129 \cdot 10^{-3}$	$1.497 \cdot 10^{-3}$	$9.779 \cdot 10^{-4}$	14345

Table 3: Application of the triangle enclosure algorithm in order to compute $\mathcal{M}_S$ for data set 2. The results are computed for several edge lengths a. In some degenerated geometric situations the boundary approximation error can be larger than the edge length of the equiangular triangles.

ε_b, δ	Computation time [s]	Boundary accuracy subset 1	subset 2	subset 3	Number of vertices
10^{-2}	0.86	$2.401 \cdot 10^{-2}$	$2.571 \cdot 10^{-2}$	$1.572 \cdot 10^{-2}$	94
10^{-3}	2.90	$1.530 \cdot 10^{-3}$	$1.804 \cdot 10^{-3}$	$1.407 \cdot 10^{-3}$	194
10^{-4}	7.19	$1.887 \cdot 10^{-4}$	$2.286 \cdot 10^{-4}$	$3.508 \cdot 10^{-4}$	404

Table 4: Application of the polygon inflation algorithm in order to compute $\mathcal{M}_S$ for data set 2. The results are computed for some settings of the control parameters ε_b (boundary precision) and δ (stopping criterion). The adaptive approximation strategy works reliably.

ε_b, δ	Computation time [s]	Boundary accuracy subset 1	subset 2	subset 3	Number of vertices
10^{-2}	3.70	$9.074 \cdot 10^{-3}$	$1.634 \cdot 10^{-2}$	$1.321 \cdot 10^{-2}$	78
10^{-3}	7.09	$1.446 \cdot 10^{-3}$	$2.654 \cdot 10^{-3}$	$1.997 \cdot 10^{-3}$	146
10^{-4}	16.72	$1.470 \cdot 10^{-4}$	$7.401 \cdot 10^{-4}$	$2.886 \cdot 10^{-4}$	30

Table 5: Application of the inverse polygon inflation algorithm in order to compute $\mathcal{M}_S$ for data set 2. The results are computed for some settings of the control parameters ε_b (boundary precision) and δ (stopping criterion). By construction, the inverse polygon inflation algorithm requires at least twice the computing time compared to the (direct) polygon inflation algorithm.

ε_b	Number of rays N_{basic}	N (final)	Computation time [s]	Boundary accuracy subset 1	subset 2	subset 3
100	286	$1 \cdot 10^{-3}$	13.46	$1.730 \cdot 10^{-3}$	$6.993 \cdot 10^{-3}$	$1.241 \cdot 10^{-3}$
100	368	$5 \cdot 10^{-4}$	20.05	$9.349 \cdot 10^{-4}$	$7.350 \cdot 10^{-4}$	$7.247 \cdot 10^{-4}$
100	786	$1 \cdot 10^{-4}$	50.82	$1.876 \cdot 10^{-4}$	$1.994 \cdot 10^{-4}$	$4.570 \cdot 10^{-4}$

Table 6: Application of the ray casting method with a simple adaptive control strategy in order to compute $\mathcal{M}_S$ for data set 2. The results are computed for some boundary precision parameters ε_b and a number of $N_{\text{basic}} = 100$ initial rays.