A ray casting method for the computation of the area of feasible solutions for multicomponent systems: Theory, applications and *FACPACK*-implementation.

Mathias Sawall[a], Klaus Neymeyr[a,b]

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

Abstract

Multivariate curve resolution methods suffer from the non-uniqueness of the solutions. The set of possible nonnegative solutions can be represented by the so-called Area of Feasible Solutions (AFS). The AFS for an s-component system is a bounded $(s-1)$-dimensional set. The numerical computation and the geometric construction of the AFS is well understood for two- and three-component systems but gets much more complicated for systems with four or even more components.

This work introduces a new and robust *ray casting method* for the computation of the AFS for general s-component systems. The algorithm shoots rays from the origin and records the intersections of these rays with the AFS. The ray casting method is computationally fast, stable with respect to noise and is able to detect the various possible shapes of the AFS sets. The easily implementable algorithm is tested for various three- and four-component data sets.

Key words: nonnegative matrix factorization, area of feasible solutions, band boundaries of feasible solutions, polygon inflation, *FACPACK*.

1. Introduction

In computer graphics ray tracing is the fundamental method for producing views of three-dimensional objects on a computer screen. A forerunner of ray tracing is the so-called "ray casting" as presented by Appel in 1968 [3]. The basic idea of ray casting is to shoot light rays from the eye of an observer to his neighborhood and to detect the nearest blocking objects in the path of these rays. These points are used to cast the visible surface of the objects. Here we extend the idea of ray casting in a way that we record the path of the ray through the object. By assembling all these in-object paths we approximate the shape of the complete object. We use this algorithm in order to approximate the shape of certain bounded sets which represent the set of all possible nonnegative matrix factorizations (NMF) of a given spectral data matrix in chemometrics. Next we introduce the chemometric problem.

The series of spectra taken from a chemical reaction system can be stored in the rows of a matrix D. Multivariate curve resolution (MCR) methods aim at extracting from these superposed multicomponent spectral data the pure component information. If the chemical reaction system contains s chemical components, then we are interested in matrix factorizations $D = CA$, where C contains in its s columns the concentration profiles (along the time axis) of the pure components and where A contains in its s rows the associated pure component spectra. Thus $D = CA$ is the matrix form of the Lambert-Beer law. The key difficulty of this matrix factorization problem is the so-called rotational ambiguity of the solution [44, 2]. This means that many factorizations $D = CA$ exist with nonnegative factors C and A. For reaction systems with two components this ambiguity was first analyzed by Lawton and Sylvestre [21]. Later Borgen and Kowalski [7] as well as Rajkó and István [32] extended this analysis to three-component systems. These authors use the Area of Feasible Solutions (AFS) in order to represent the set of all nonnegative/feasible solutions in a low-dimensional way. Effective numerical algorithms for the approximate computation of the AFS have been devised by Golshan, Abdollahi and Maeder [12] as well as Sawall, Neymeyr et al. [37, 38]. An alternative approach is the direct computation of minimal and maximal band boundaries as suggested by Gemperline [9] and Tauler [43]. See also [11] for a review on recent AFS methods.

A strong advantage of the AFS approach is that it makes available *all* NMFs. One might call this a *global*

approach. Hence the AFS comprises all the solutions which can be put out by any MCR method. It is a well-known fact that different MCR methods may result in very different results. Having computed the AFS for a given spectral data matrix, various post-processing techniques can help to eliminate chemically non-relevant solutions in the AFS. This reduction can be done in a controlled and steerable way. To this end, additional soft constraints can be used to reduce the AFS, see [4, 41, 31]. Alternatively, computational techniques like the window factor analysis [26], the evolving factor analysis [24, 22] or kinetic modeling [8, 33] can be employed. See also Malinowski [25], Hamilton and Gemperline [17] as well as Maeder and Neuhold [23] for a general and deepened introduction to MCR.

The aim of this paper is to present ray casting as a robust and fast numerical method for the computation of the AFS. In principle, ray casting can be applied to chemical reaction systems with any number of components. If the spectral data matrix D belongs to a chemical reaction system which includes a number of s chemical components and D is a rank-s matrix, then the AFS is a bounded $(s-1)$-dimensional set. The ray casting algorithm records the paths of "equiangular" rays starting at the origin on its way through the AFS. All these paths serve to approximate the AFS.

In particular, we demonstrate applications to three- and four-component systems. For four-component systems the ray casting method has been implemented in a new module of the *FACPACK* software [39]. In a further new software module the complementarity and coupling theory [29, 36, 40, 34] has been implemented for four-component systems.

Organization of the paper: Section 2 briefly introduces the AFS concept and the fundamentals of its numerical approximation. In Section 3 important properties of the AFS are explained and the new gap-free intersection property is introduced and proved. The ray casting method is explained and discussed in Section 4. Section 5 presents the background of its *FACPACK* implementation. Finally, numerical results for three- and four-component systems are presented in Section 6.

2. The AFS

2.1. Curve resolution methods

The basic equation for the two-way curve resolution problem is the Lambert-Beer law. Its matrix formulation in absence of noise reads

$$D = CA$$

with the spectral (mixture) data matrix $D \in \mathbb{R}^{k\times n}$ and with $C \in \mathbb{R}^{k\times s}$ and $A \in \mathbb{R}^{s\times n}$. The number k stands for the number of measured spectra and n is the number of spectral channels of each spectrum. Further, s denotes the number of independent components with $s \le \min(k, n)$. Our problem is to extract from given D the "true" chemically meaningful pure component decomposition $D = CA$. As already mentioned in the introduction, the rotational ambiguity makes this factorization problem hard to solve. It is an ill-posed problem.

The chemometric literature contains many techniques and algorithms to compute more or less suitable factorizations. Sometimes only parts of the factors are computed. All this reduces the rotational ambiguity. Among many others, some of these MCR approaches are hard-modeling techniques [16, 8], window factor analysis [24, 22, 26] in combination with uniqueness theorems [27], complementarity and coupling theorems [34] and soft modeling [44, 30, 13].

Instead of focusing on a single factorization, one can take up the challenge to compute the set of *all* NMFs of D. We call such factorizations with $C \ge 0$, $A \ge 0$ and $D = CA$ feasible solutions. In the following, we describe techniques to get the set of all feasible spectral factors $A \ge 0$. A set of matrices is difficult to handle. However, the AFS provides an approach for a low-dimensional representation of these matrices. The key idea is to consider only the expansion coefficients of the possible spectra with respect to a basis of right singular vectors of the matrix D. The AFS is introduced in Section 2.3; see also [7, 32, 12, 37, 38, 19, 35].

2.2. SVD-based factorization

The most usual way to solve the factorization problem $D = CA$ is to start with a singular value decomposition (SVD) of $D \in \mathbb{R}^{k\times n}$. The SVD has the form $D = U\Sigma V^T$. For an s-component system and if D has the rank s, then s singular values are larger than zero. These nonzero singular values are the diagonal elements of the diagonal matrix $\Sigma \in \mathbb{R}^{s\times s}$. The columns of the orthogonal matrices $U \in \mathbb{R}^{k\times s}$ and $V \in \mathbb{R}^{n\times s}$ are the left and right singular vectors. The concentration profiles of the pure components are presented by a linear expansion in terms of the left singular vectors. The pure component spectra are presented by expansions in terms of the right singular vectors. These representations of C and A can be expressed by an invertible matrix $T \in \mathbb{R}^{s\times s}$ (whose matrix elements are the expansion coefficients for A) so that

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

see [21]. The expansion coefficient of the first right singular vector and the left singular vector are always nonzero as otherwise A or C would have negative entries; see Theorem 2.2 in [38] for the mathematical justification by the Perron-Frobenius theory on properties of eigenvalues and eigenvectors of nonnegative matrices. This together with a scaling argument justifies that the first column of T can be filled with ones, i.e.

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

2.3. The AFS

A permutation can be applied to the rows of T in (1). The inverse permutation must be applied simultaneously to the columns of T^{-1}. Thus the set of all possible nonnegative spectra is completely represented by the set of all possible *first rows* of T. This is the basic idea behind the AFS construction. Therefore the AFS consists of all row vectors $x \in \mathbb{R}^{1\times(s-1)}$ so that

$$T = \begin{pmatrix} 1 & x_1 & \cdots & x_{s-1} \\ 1 & & & \\ \vdots & & S & \\ 1 & & & \end{pmatrix} \tag{2}$$

in (1) leads to feasible factors $C = U\Sigma T^{-1} \geq 0$ and $A = TV^T \geq 0$. Therein S is a proper $(s-1)\times(s-1)$ matrix which is to be determined, e.g., by an optimization. We denote the AFS by $\mathcal{M}$. Thus $\mathcal{M}$ has the form

$$\mathcal{M} = \{x \in \mathbb{R}^{1\times(s-1)} : T \text{ in (2) is a regular matrix}, \tag{3}$$
$$C = U\Sigma T^{-1} \geq 0 \text{ and } A = TV^T \geq 0\}. \tag{4}$$

The analysis and geometric construction of the AFS was first studied for two-component systems in [21]. In [7, 32, 19] the AFS is geometrically constructed for $s = 3$ components. Numerical methods for the AFS-approximation for three-component systems are the brute-force grid search approach [2], the triangle-boundary-enclosure algorithm [12] and the polygon-inflation method [37, 38]. The paper [10] contains a recent application of polygon-inflation. For $s = 4$, up to now only the sliced triangle-boundary-enclosure algorithm has been introduced, see [14]. The algorithmic idea of polygon-inflation [37] can also be extended to a polyhedron inflation for four-component systems.

For the numerical approximation of the AFS the crucial step is to decide whether a certain point x is feasible or not. This step is also the computationally most time-consuming part of the algorithm due to its repeated application. For the feasibility decision we use the target function $f : \mathbb{R}^{(s-1)} \times \mathbb{R}^{(s-1)\times(s-1)} \to \mathbb{R}$ defined as

$$\begin{aligned} f(x,S) = &\sum_{i=1}^{s}\sum_{j=1}^{k} \min(0, C_{ji})^2 \\ &+ \sum_{i=1}^{s}\sum_{j=1}^{n} \min(0, A_{ij})^2 + \|I_s - TT^+\|_F^2. \end{aligned} \tag{5}$$

With this f the AFS has the form

$$\mathcal{M} = \{x \in \mathbb{R}^{1\times(s-1)} : \min_{S\in\mathbb{R}^{(s-1)\times(s-1)}} f(x,S) = 0\}.$$

For data including noise, nonlinearities or perturbations the function f needs to be modified, see for example f from [37].

In the following we also need an important superset of $\mathcal{M}$ which contains all x so that the linear combinations $(1,\ x)\cdot V^T$ are componentwise nonnegative

$$\mathcal{M}^+ = \{x \in \mathbb{R}^{1\times(s-1)} : (1,\ x)\cdot V^T \geq 0\}. \tag{6}$$

The set $\mathcal{M}^+$ is called FIRPOL by Borgen and Kowalski [7]. Geometrically FIRPOL is a polygon which specifies the outer boundary of the set of the feasible vectors x.

3. The gap-free intersection property of the AFS

This section explains the three important properties which constitute the basis for the new ray casting method. These properties are:

1. *Boundedness*: The AFS is a bounded set.
2. *Exclusion of the origin*: The origin of the AFS for an s-component system, namely the point $0 \in \mathbb{R}^{s-1}$, is never an element of the AFS $\mathcal{M}$. However, 0 is always contained in $\mathcal{M}^+$ (FIRPOL), see Eq. (6).
3. *Gap-free ray intersection*: The intersection of the AFS with a ray starting at the origin is empty or a line segment. The line segment may be degenerated to a single point. In other words, this intersection is *gap-free*.

The first two properties are not new and are recapitulated in Section 3.1. The new ray intersection property is proved in Section 3.2 by Theorem 3.3.

3.1. Boundedness of the AFS and position of the origin

The boundedness of the AFS is a necessary prerequisite for all numerical algorithms to compute the AFS. Theorem 2.3 in [38] shows that the AFS is bounded if and only if the matrix $D^T D$ is irreducible. This is a mild

condition which can be assumed to hold whenever the spectral data, loosely speaking, does not allow to split the data matrices into independent reaction subsystems.

The second property claims that the origin is an element of $\mathcal{M}^+$ (FIRPOL), but is never an element of $\mathcal{M}$ (AFS). Since the first right singular vector v_1 can be assumed as a componentwise nonnegative vector by the Perron-Frobenius theory, see Lemma 3.1, it holds that

$$(1, \underbrace{0, \ldots, 0}_{s-1\text{-times}})V^T = v_1 \geq 0.$$

Thus $x = 0$ is an element of $\mathcal{M}^+$ according to Equation (6). Further, Theorem 2.5 in [38] proves that the origin is not an element of $\mathcal{M}$.

3.2. The gap-free intersection property

This section presents and proves the ray intersection property of the AFS, which is the basis for the ray casting method. The idea can easily be explained: If a ray is shot from the origin and if the ray hits the surface of the AFS at a certain point x of the AFS $\mathcal{M}$, then the intersection of this ray with the AFS consists of all points γx with $1 \leq \gamma \leq \gamma^*$ so that $\gamma^* x$ belongs to the surface of the set FIRPOL $\mathcal{M}^+$, see Figure 1. In other words a nonempty intersection of such a ray with the AFS is a line segment (which may degenerate to a single point). This intersection property is proved by Theorem 3.3. First, two preparatory lemmata are needed.

Lemma 3.1. *Let $D \in \mathbb{R}^{k\times n}$ be a nonzero nonnegative matrix. Let $U\Sigma V^T$ be a singular value decomposition of D. If the first singular vector $V(:,1)$ is a componentwise nonnegative vector, then the first left singular vector $U(:,1)$ is also a componentwise nonnegative vector. (The Perron-Frobenius theory [28] guarantees that nonnegativity of $V(:,1)$ can always be attained by a proper choice of this singular vector.)*

Proof. The equation $DV = U\Sigma$ implies that

$$DV(:,1) = U(:,1)\sigma_1$$

with $\sigma_1 > 0$ since $D \neq 0$. Thus $V(:,1) \geq 0$ and $D \geq 0$ show that $DV(:,1) \geq 0$. Hence $U(:,1) \geq 0$. □

Lemma 3.2. *Let $C \in \mathbb{R}^{k\times s}$ and $A \in \mathbb{R}^{s\times n}$ be nonnegative matrices. Further let*

$$X = \begin{pmatrix} \beta & b^T \\ 0 & I \end{pmatrix} \in \mathbb{R}^{s\times s} \tag{7}$$

with $\beta \geq 1$ and $b \in \mathbb{R}^{s-1}$ with $b \leq 0$ componentwise. Further, I denotes the $(s-1)\times(s-1)$ identity matrix. Then

$$\widetilde{C} = CX^{-1} \tag{8}$$

is a nonnegative matrix. Moreover $X(1,:)A \geq 0$ implies that

$$\widetilde{A} = XA \tag{9}$$

is also a nonnegative matrix.

Proof. Direct computation shows that

$$X^{-1} = \begin{pmatrix} 1/\beta & -b^T/\beta \\ 0 & I \end{pmatrix}.$$

The assumptions on β and b imply that X^{-1} is a nonnegative matrix. Thus $\widetilde{C} = CX^{-1}$ is also nonnegative. By assumption $\widetilde{A}(1,:) = X(1,:)A$ is nonnegative. For the remaining rows of $\widetilde{A}$ the equations (7) and (9) imply that

$$\widetilde{A}(j,:) = A(j,:) \geq 0 \quad \text{for } j = 2, \ldots, s.$$

This proves $\widetilde{A} \geq 0$. □

Lemmata 3.1 and 3.2 together with Corollary 2.3 from [38] help to prove that nonempty intersections of the AFS with rays starting at the origin are line segments. A line segment may even be degenerated to a single point.

Theorem 3.3 (On the intersections of rays with the AFS)**.** *Let $D \in \mathbb{R}^{k\times n}$ be a nonnegative matrix with no zero-column, no zero-row and* rank$(D) = s$ *so that $D^T D$ and DD^T are irreducible matrices. Let $U\Sigma V^T$ be a truncated SVD of D with $V(:,1) > 0$.*

If $x \in \mathcal{M}$, then a number $\gamma^* \geq 1$ exists so that $\gamma^* x$ is located on the boundary of $\mathcal{M}^+$ (FIRPOL) and the line segment

$$\{\gamma x : \ \gamma \in [1, \gamma^*]\}$$

is a subset of the AFS $\mathcal{M}$.

Proof. The feasible point $x \in \mathcal{M}$ belongs to a matrix T of the form (with the all-ones vector $e = (1, \ldots, 1)^T \in \mathbb{R}^{s-1}$)

$$T = \begin{pmatrix} 1 & x^T \\ e & S \end{pmatrix}$$

so that $C = U\Sigma T^{-1}$ and $A = TV^T$ are nonnegative matrices. Our aim is to show that $\widetilde{C} = U\Sigma\widetilde{T}^{-1}$ and $\widetilde{A} = \widetilde{T}V^T$ are also nonnegative with

$$\widetilde{T} = \begin{pmatrix} 1 & \gamma x^T \\ e & S \end{pmatrix}$$

for any $\gamma \in [1, \gamma^*]$. Thus with $X = \widetilde{T}T^{-1}$ we have

$$\widetilde{C} = \underbrace{(U\Sigma T^{-1})}_{C}\underbrace{(T\widetilde{T}^{-1})}_{X^{-1}} \quad \text{and} \quad \widetilde{A} = \underbrace{(\widetilde{T}T^{-1})}_{X}\underbrace{(TV^T)}_{A}$$

which is a representation of $\widetilde{C}$ and $\widetilde{A}$ according to (8) and (9). In the following we show that X satisfies the assumptions of Lemma 3.2. Then Lemma 3.2 proves the desired nonnegativity of $\widetilde{C}$ and $\widetilde{A}$. The inverse of the 2×2 block matrix T reads

$$T^{-1} = \begin{pmatrix} 1 + x^T\mathcal{Z}^{-1}e & -x^T\mathcal{Z}^{-1} \\ -\mathcal{Z}^{-1}e & \mathcal{Z}^{-1} \end{pmatrix} \tag{10}$$

with the Schur complement $\mathcal{Z} = S - ex^T$, see [15]. (One can easily check that $TT^{-1} = I$ holds.) Lemma 2.3 in [38] says that a componentwise nonnegative vector $U\Sigma w$ necessarily requires that the first component w_1 is nonzero. Under our assumptions on D and the SVD, Lemma 3.1 says that $U(:, 1) \geq 0$. Hence $w_1 > 0$ must hold as otherwise for $w_1 < 0$ the vector $U\Sigma w$ necessarily contains negative components. This argumentation applied to each column of $C = U\Sigma T^{-1}$ implies that the first row of T^{-1} by (10) is strictly positive. In other words we have that

$$x^T\mathcal{Z}^{-1} < 0 \tag{11}$$

holds. The sum of the components of $x^T\mathcal{Z}^{-1}$ equals $x^T\mathcal{Z}^{-1}e$. We conclude from (11) that

$$x^T\mathcal{Z}^{-1}e < 0. \tag{12}$$

The two inequalities (11) and (12) are used in the final part of the proof.

The Sherman-Morrison formula [15] allows to write the inverse Schur complement explicitly

$$\mathcal{Z}^{-1} = S^{-1} + \frac{S^{-1}ex^TS^{-1}}{1 - x^TS^{-1}e}.$$

Direct computation shows that

$$X = \begin{pmatrix} 1 - (\gamma - 1)x^T\mathcal{Z}^{-1}e & (\gamma - 1)x^T\mathcal{Z}^{-1} \\ 0 & I \end{pmatrix}.$$

As $\gamma \geq 1$, we see that the vector b in (7) equals

$$b = (\gamma - 1)x^T\mathcal{Z}^{-1}$$

and that $b \leq 0$ componentwise due to (11). Thus the first assumption of Lemma 3.2 is fulfilled. Further, $\beta = 1 - (\gamma - 1)x^T\mathcal{Z}^{-1}e \geq 1$ since $x^T\mathcal{Z}^{-1}e < 0$ by Eq. (12) and $\gamma \geq 1$. Thus the second assumption of Lemma 3.2 is fulfilled. So Lemma 3.2 guarantees that $\widetilde{C} = CX^{-1}$ is a componentwise nonnegative matrix.

Finally, we have to show that $\widetilde{A} = XA$ is a componentwise nonnegative matrix. Lemma 3.2 says that $X(1, :)A \geq 0$ is sufficient for the nonnegativity of $\widetilde{A}$. It holds that

$$X(1, :)A = a_1^T - (\gamma - 1)\underbrace{\left[x^T\mathcal{Z}^{-1}ea_1^T - x^T\mathcal{Z}^{-1}A([2 : s], :)\right]}_{=:r}$$

where a_1^T denotes the first row of A. For $\gamma = 1$ the spectrum a_1^T is feasible by assumption, i.e. $a_1^T \geq 0$. Further, for $\gamma = \gamma*$

$$X(1, :)A = a_1^T - (\gamma^* - 1)r$$

is also feasible by assumption so that $a_1^T - (\gamma^* - 1)r \geq 0$. Finally, $\gamma^* > 1$ can be assumed. Multiplication of the last inequality by the nonnegative constant $(\gamma - 1)/(\gamma^* - 1)$ with $\gamma \geq 1$ results in

$$(\gamma - 1)/(\gamma^* - 1)a_1^T - (\gamma - 1)r \geq 0.$$

As $\omega := 1 - ((\gamma - 1)/(\gamma^* - 1)) \geq 0$ for $1 \leq \gamma \leq \gamma^*$ we can add $\omega a_1^T \geq 0$ to the last inequality which results in

$$a_1^T - (\gamma - 1)r \geq 0$$

which shows that $X(1, :)A \geq 0$. Lemma 3.2 proves $\widetilde{A} \geq 0$. □

Theorem 3.3 implies Corollary 3.4.

Corollary 3.4. *On the assumptions of Theorem 3.3 let x be a point on the boundary of $\mathcal{M}^+$ (FIRPOL). If x is not feasible, then the ray which starts at the origin and which runs through x contains no feasible point.*

Proof. Assume the existence of a point $y = \omega x \in \mathcal{M}$ with $0 < \omega < 1$. Then Theorem 3.3 says that the line segment from y to x is contained in the AFS. This contradicts the assumption that x is not in the AFS. □

4. The ray casting method

This section provides a detailed explanation of the new ray casting algorithm for approximating the AFS. The single steps of the algorithm are fundamentally based on the three properties of the AFS, namely the

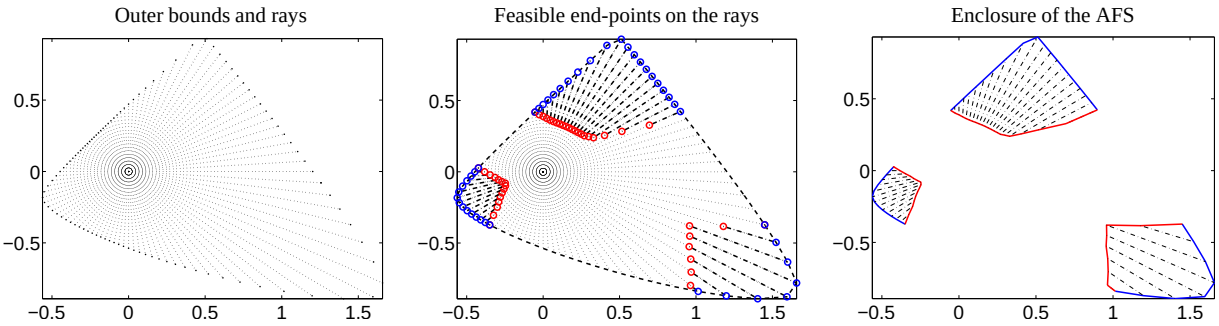

Figure 1: Construction of the 2D AFS for a three-component system by $m = 100$ rays. Left: Equiangular ray casting with ray directions $v_i \in \mathbb{R}^2$ (step 1). Computation of outer bounds radii R_i (step 2). Center: Computation of the inner radii r_i (step 3). The points of the set $\mathcal{R}^{\text{out}}$ are marked by blue circles and the points of the set $\mathcal{R}^{\text{in}}$ are marked by red circles. Right: Connection of the inner and outer points (step 4). The boundary sections on the boundary of FIRPOL are colored in blue. Boundary sections on the inner boundary are colored in red. Also, the connecting lines between inner and outer boundaries are colored in red. For demonstration purposes the angle-resolution is very low. Typically much more than $m = 100$ rays are used.

boundedness, the *exclusion of the origin* and the *gap-free ray intersection*, see Section 3. The ray casting algorithm works with equiangular rays (e.g. in polar coordinates) which start at the origin. The origin is never a feasible point. If a certain point x is feasible, then the complete line segment of this ray from x to the boundary of $\mathcal{M}^+$ (FIRPOL) belongs to the AFS. Conversely, if the intersection of the ray with the boundary of $\mathcal{M}^+$ is not a feasible point, then the intersection of this ray with the AFS is empty. Computationally, we first check the feasibility of this point of intersection. In a second step, we look for the point on this ray which is feasible and closest to the origin. In a final step, all these extremal points are connected in order to approximate the boundary of the AFS.

4.1. Notation and algorithm

We consider a number of m vectors v_i, $i = 1, \ldots, m$, starting at the origin of the AFS. An exemplary, equiangular 2D setting of these vectors is shown in Figure 1 (left subplot) and a 3D setting is shown in Figure 2 (left subplot). For ease of presentation we do not distinguish the vectors v_i from the rays $\{cv_i : c \geq 0\}$. If the ray v_i hits the AFS, then two radii are determined. These are the minimal radius r_i, i.e. the distance of the closest AFS-point on v_i to the origin, and the maximal radius R_i, which is the maximal distance of an AFS-point on this ray to the origin. Mathematically, these radii for a given ray v_i are

$$r_i = \min\left\{\gamma > 0 \text{ with } \min_S f(\gamma v_i, S) = 0\right\}, \qquad (13)$$

$$R_i = \max\left\{\gamma > 0 \text{ with } \min_S f(\gamma v_i, S) = 0\right\} \qquad (14)$$

with $f(x, S)$ by (5). Hence R_i is the distance of the point of intersection of the ray v_i with the boundary of $\mathcal{M}^+$ (FIRPOL). Theorem 3.3 proves that the line segment

$$\{\gamma v_i : \ r_i \leq \gamma \leq R_i\} \qquad (15)$$

equals the intersection of the ray v_i with the AFS. We denote by $\mathcal{R}^{\text{in}}$ the set of all points with minimal radii on the rays v_i, $i = 1, \ldots, m$, and by $\mathcal{R}^{\text{out}}$ the corresponding set of points with maximal radii.

The steps of the ray casting algorithm are as follows:

1. Assign a number of m equiangular rays v_i in the $(s-1)$-dimensional space, e.g., by using (generalized) polar coordinates.
2. For each ray compute the maximal radius R_i so that $R_i v_i$ is located on the surface of $\mathcal{M}^+$ (FIRPOL).
3. For each ray test whether $R_i v_i \in \mathcal{M}$ or not. If $R_i v_i \in \mathcal{M}$, then compute the minimal radius $r_i = \min\{\gamma > 0 : \ \gamma v_i \in \mathcal{M}\}$.
4. Connect the matching interior points $r_i v_i$ and associated exterior points $R_i v_i$ in order to approximate the boundary of (a segment) of the AFS.

4.2. Computation of the extremal points

For noise-free data the computation of the maximal radius R_i can either be done by direct evaluation of the nonnegativity constraints or alternatively by means of numerical bisection. The computational costs are low as only the FIRPOL constraint $(1, x)V^T \geq 0$ is to be checked along the ray. The computation of the radius r_i is done by using the bisection method along each ray. We use the bisection method with an additional control parameter ε_b in a way that

$$\min_S f(r_i v_i, S) = 0 \ \text{ and } \ \min_S f((r_i - \varepsilon_b)v_i, S) > 0 \qquad (16)$$

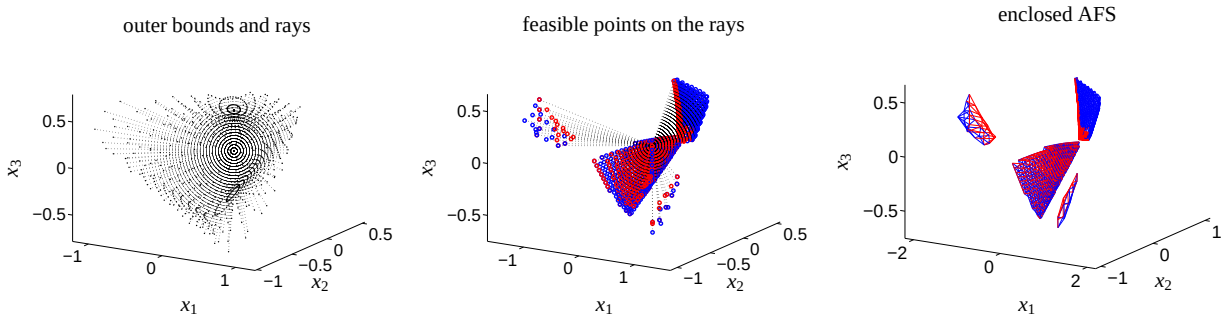

Figure 2: Computation of a three-dimensional AFS for a four-component system ($s = 4$) with $m = 512$ rays. Left: 512 rays $v_i \in \mathbb{R}^3$ (step 1) are shown with end-points on the boundary of FIRPOL. The length of the ray v_i is R_i (step 2). Center: Computation of the inner radii r_i (step 3). Only feasible directions are plotted. Right: Matching points on the surfaces of the segments of the AFS are connected by grid lines (step 4). As in Figure 1 the points and lines on the outer boundary are drawn in blue and those on the inner boundary are drawn in red.

with f by (5). In each step a nonlinear optimization problem on finding a proper S must be solved. Thus the computational costs are relatively high.

4.3. Precision, resolution and corner cut-off

The numerical accuracy of the radii r_i and R_i can be controlled by the parameter ε_b in (16). The density of the extremal points $\mathcal{R}^{\text{in}}$ and $\mathcal{R}^{\text{out}}$ on the surface of the AFS is determined by the number of rays m. A large number m of rays results in a fine lateral resolution of the boundary of the AFS.

A critical aspect of the ray casting method is that a low angle-resolution, i.e. m is small, may result in a corner cut-off. Especially if a first ray has a relatively large intersection with the AFS and a second adjacent ray contains no feasible point, then some part of the boundary of the AFS approximation is determined by the first ray. The cone between these two rays may include a part of the AFS which is then cut off. An increased resolution by a larger number m of rays can reduce such a cut-off of boundary-near regions of the AFS. Figure 3 demonstrates a minor corner cut-off for a two-dimensional AFS.

4.4. Two- and three-component systems

The computation of the AFS for two- and three-component systems ($s = 2$ or $s = 3$) is well-understood, see for example [21, 7, 32, 1, 2, 12, 37, 38].

For two-component systems ($s = 2$) the AFS consists of two separated intervals, i.e.

$$\mathcal{M} = [-R_1, -r_1] \cup [r_2, R_2]$$

with the sets of positive minimal and maximal radii $\mathcal{R}^{\text{in}}$ and $\mathcal{R}^{\text{out}}$, see Section 4.1,

$$\mathcal{R}^{\text{in}} = \{r_1, r_2\}, \quad \mathcal{R}^{\text{out}} = \{R_1, R_2\}.$$

Computationally, these radii, which represent the boundaries of $\mathcal{M}$, can be determined by starting the algorithm with (the only possible) $m = 2$ rays given by $v_1 = -1$ and $v_2 = 1$.

For three-component systems ($s = 2$) the AFS is a subset of the plane $\mathbb{R}^2$. A number of m equiangular rays is constructed in polar coordinates

$$v_i = \begin{pmatrix} \cos\phi_i \\ \sin\phi_i \end{pmatrix} \quad \text{with} \quad \phi_i = 2\pi\frac{i-1}{m}, \quad i = 1, \ldots, m. \tag{17}$$

Figure 1 shows the application to the model problem of a three-component consecutive reaction with shifted Gauss profile spectra taken from Section 2.1 of [39]. A number of $m = 100$ rays are used for the approximation of the boundary of the AFS which consists of three isolated subsets, which we call the segments of the AFS.

4.5. Four components

For four-component systems ($s = 4$) the AFS is a three-dimensional set. Spherical coordinates can be used to construct evenly distributed (and with respect to $\phi_{i,1}$ and $\phi_{j,2}$ equiangular) rays

$$v_{i+j} = \begin{pmatrix} \sin(\phi_{i,1})\cos(\phi_{j,2}) \\ \sin(\phi_{i,1})\sin(\phi_{j,2}) \\ \cos(\phi_{i,1}) \end{pmatrix}, \quad \phi_{i,1} = \pi\frac{i-1}{\ell_1}, \ \phi_{j,2} = 2\pi\frac{j-1}{\ell_2}, \tag{18}$$

for $i = 1, \ldots, \ell_1$ and $j = 1, \ldots, \ell_2$. This allows to form $m = \ell_1\ell_2$ rays. Figure 2 shows the application to a model problem ($D \in \mathbb{R}^{10\times 11}$ with C formed by Gaussians and A formed by partially overlapping and shifted Gaussians) with $\ell_1 = 13$ and $\ell_2 = 32$. However, the resolution with $m = 512$ rays is relatively low and is only used for demonstration purposes.

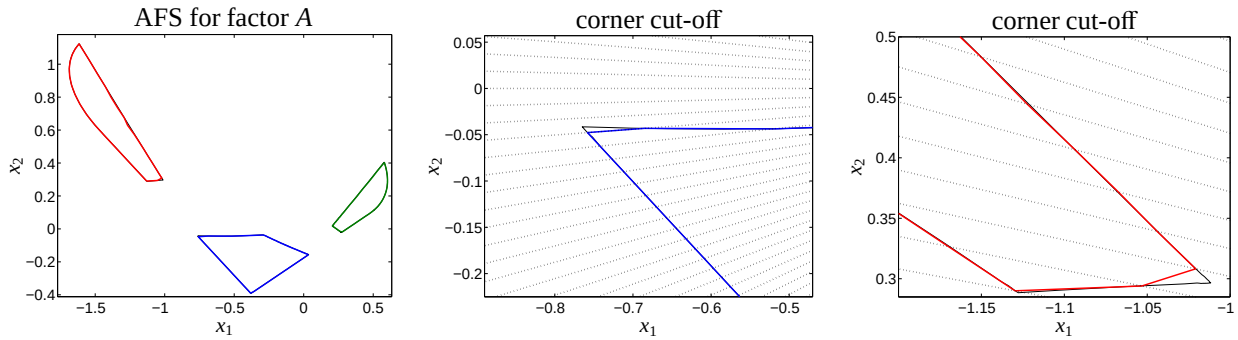

Figure 3: AFS approximations for the model problem from Sec. 6.1. Left: The colored boundary curves of the AFS have been computed by ray casting with $m = 300$ rays with a boundary precision of $\varepsilon_b = 10^{-3}$. The (for the most part underlying) black solid lines result from the polygon inflation method by the *FACPACK* software [38]. Center and right: Enlargements of the left-upper area of the blue AFS segment and the right lower area of the red segment. The dotted lines are the rays of the ray casting algorithm. A minor corner cut-off can be stated, see Sec. 4.3.

4.6. Higher number of components

For s-component systems with $s \geq 5$ the rays $v_i \in \mathbb{R}^{s-1}$ can be constructed by using $(s-1)$-dimensional spherical coordinates [6]

$$\begin{aligned} v_{i,1} &= \cos(\phi_{i,1}), \\ v_{i,2} &= \sin(\phi_{i,1})\cos(\phi_{i,2}), \\ v_{i,3} &= \sin(\phi_{i,1})\sin(\phi_{i,2})\cos(\phi_{i,3}), \\ &\vdots \\ v_{i,s-2} &= \sin(\phi_{i,1})\cdot\ldots\cdot\sin(\phi_{i,s-3})\cos(\phi_{i,s-2}), \\ v_{i,s-1} &= \sin(\phi_{i,1})\cdot\ldots\cdot\sin(\phi_{i,s-3})\sin(\phi_{i,s-2}). \end{aligned} \tag{19}$$

The discretization for the last angle $\phi_{i,s-2}$ is

$$\phi_{i,s-2} = 2\pi\frac{i-1}{\ell_{s-2}}, \qquad i = 1,\ldots,\ell_{s-2},$$

and for the remaining angles we use

$$\phi_j = \pi\frac{i-1}{\ell_j}, \qquad i = 1,\ldots,\ell_j, \quad j = 1,\ldots,s-3.$$

This general definition already includes the cases of polar coordinates ($s = 3$) and 3D spherical coordinates ($s = 4$). The total number of rays is $m = \ell_1\ell_2\ldots\ell_{s-2}$.

4.7. Stability for the various shapes of an AFS

In general, the AFS can be a topologically connected set with holes, or the AFS can consist of several isolated subsets, the segments of an AFS. Especially for two-component systems the AFS consists of two separated intervals, one only in the negative numbers and the other only in the positive numbers. For $s = 3$ the AFS is well known to be either a single-segment AFS with a hole around the origin [37, 38] or it consists of three isolated segments. Additionally, nonnegative matrices can be constructed whose AFS consists of 6 segments, 9 segments or even higher multiples of 3 segments. However, an AFS for a real chemical reaction system with more than 3 segments is not known up to now. For $s = 4$ there is no general presumption on the possible number of segments. Figure 7 shows an AFS with four clearly separated segments. Further Figure 9 presents a single-segment AFS with a complicated "Swiss-cheese like" hole structure. Further experiments have shown that the AFS for the case $s = 4$ may also have a closed surface with no holes going through to the origin. However, the origin is never contained in the AFS - any AFS has a hole around the origin.

The ray casting algorithm is general and flexible enough to compute approximations of all AFS types. The gap-free intersection property of the AFS by Theorem 3.3 guarantees that the ray casting algorithm can work successfully for any number s of components. This is a main advantage compared to the triangle-boundary-enclosure algorithm [12, 14] or the basic polygon inflation method [37]. However, the inverse polygon inflation algorithm [38] and the geometric constructive Borgen plot approach [7, 32, 19] also work for any AFS of a three-component system.

4.8. Factor-locking

Sometimes certain pure component spectra or concentration profiles are known, e.g., a spectrum of the main reactant or main product. This additional information reduces the rotational ambiguity [34, 29]. If particular columns of C or rows of A are available, then the equation $D = CA$ provides some information on the other factor (A or C). See, for example, [5, 40, 29] for a systematic analysis of these mutual restrictions for systems with $s = 3$ or $s = 4$ chemical components. Given

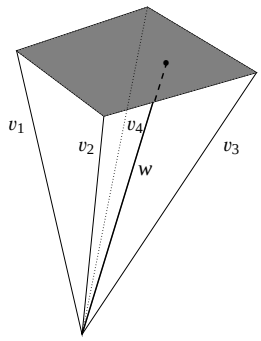

Figure 4: Detection of degenerated AFS segments: The initial ray w (computed by an NMF) has a feasible point on the surface of $\mathcal{M}^+$ (FIRPOL). The gray rectangle is a part of the surface of $\mathcal{M}^+$. The pyramid spanned by the closest neighboring rays v_i, $i = 1, \ldots, 4$, serves to check the degree of degeneracy of the AFS segment whether it is a line segment or plane segment.

components are represented in the AFS by fixed points. This type of additional information is called an equality constraint [5]; in the *FACPACK* software [38] equality constraints can be set by factor-locking. Such a constraint also reduces the size of the AFS segments for the remaining unknown factors.

Remark 4.1. *Let one or more points in the AFS be given, i.e. certain pure component spectra or concentration profiles are fixed. Then the gap-free intersection property (see Theorem 3.3) still holds for the reduced AFS-segments of the remaining components.*

The proof of Remark 4.1 is very similar to the proof of Theorem 3.3. Once again, one can prove that the submatrix S of T which represents a certain vector x of the AFS can also be used to represent the stretched point γx. Locked points of the AFS are represented by certain components of S which remain unchanged. Hence, γx is feasible also in the presence of these locked points. Essentially, Remark 4.1 says that the ray casting algorithm can be applied without changes to the case of active equality constraints. The resulting AFS is a subset of the original AFS (only under nonnegativity constraints).

5. Software implementation in FACPACK

The *FACPACK*-toolbox is a collection of several SMCR methods with the focus on global AFS methods [39]. The software contains implementations of the polygon inflation method and its variants [37, 38], of the geometric-constructive Borgen-plots [7, 32, 19], of the complementarity theorem [40] and its restrictions on the AFS [40] and so on. *FACPACK* is equipped with a graphical user interface in MatLab. The computational cores of the program are written in C. The toolbox can be downloaded from

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

The latest revision of the software includes a module for computing the three-dimensional AFS for ($s = 4$)-component systems by means of ray casting. Additionally, a further new module implements the complementarity theory [34] for four-component systems by means of ray casting.

5.1. Refinement process

The minimal and maximal radii r_i and R_i, see Eqns. (13) and (14), are computed for each ray by solving nonlinear optimization problems. The numerical results of these optimizations, especially for the inner points, depend on good initial guesses. In order to identify and to rule out wrongly classified points, a refinement post-process is started at the end of the AFS computation. The reliability of each minimal radius r_i is checked by restarting the optimization process with various initial values. These initial values are taken from the already computed inner radii of the closest neighboring rays.

5.2. Approximation of degenerated AFS segments

A special challenge for any numerical approximation of the AFS are cases in which the s-dimensional AFS contains lower dimensional segments.

An important example is a three-dimensional AFS for a four-component system which contains planar, linear or punctiform segments. The new 3D-AFS module of *FACPACK* can compute all these types of AFS segments. Therefore the algorithm does not only compute the minimal and maximal radii for the m rays, but also uses the initial NMF of the spectral data matrix D. For a four-component system, the initial NMF provides four feasible points and these define the directions of four

initial rays. If along one such ray the minimal radius and the maximal radius coincide, then this indicates a degenerated AFS segment. In such a case the algorithm checks the four closest neighboring rays; see Figure 4 for an illustration of the pyramid spanned by the neighboring rays together with the enclosed initial ray from the NMF. If the minimal radius and the maximal radius coincide for the initial NMF ray and if for all four neighboring rays the intersections with the AFS are either empty or consist of only a single point, then the AFS segment is degenerated. The algorithm has to distinguish planar segments from linear segments and from punctiform segments:

Planar segments:
If a certain AFS segment is not a volume segment, then the algorithm tests this segment for planarity. Any planar segment must be a subset of a plane which has been used for the construction of the surface of the polyhedron $\mathcal{M}^+$ (FIRPOL). If an initial ray contains only one feasible point (necessarily this point is located on the surface of $\mathcal{M}^+$), then an adapted polygon inflation method [37] is used in order to compute the boundary of the associated planar AFS segment, which is again located on the surface of $\mathcal{M}^+$.

Linear segments:
If a segment is neither a volume segment nor a planar segment, then the algorithm tests this segment for linearity. In order to detect a line segment, we follow the optimization strategy on the search direction as described in Sec. 4.6 of [38]. In contrast to [38] we need two different angle variables in the optimization for the computation of a linear segment in 3D. Starting from a feasible point $x \in \mathbb{R}^3$ (which might be gained by an initial NMF) the two angles ϕ_1 and ϕ_2 are determined in a way so that the point

$$x + r \begin{pmatrix} \sin\phi_1 \cos\phi_2 \\ \sin\phi_1 \sin\phi_2 \\ \cos\phi_1 \end{pmatrix} \tag{20}$$

is feasible for a sufficiently small nonzero radius r. Next the minimal and maximal bounds $r_l \leq 0$ and $r_r \geq 0$ on this line segment are computed so that the points on the line segment with $r_l \leq r \leq r_r$ in (20) are feasible. The result is a feasible line segment.

Punctiform segments:
If the segment is not a volume segment, a planar segment or a linear segment, then it must be an isolated point. A unique pure component has been found.

Figure 5 shows the AFS for the upper triangular matrix

$$D = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Its AFS consists of a unique point, a linear segment, a planar segment and a volume segment. This simple rank-4 matrix comprises all difficulties of an AFS computation for segments with different dimensionalities in an elegant and easy way. [1] This artificial example can principally be extended to a data matrix which corresponds to a chemically interpretable data set.

The unique point is computationally accessible from any initial NMF of D. The computation of the linear segment by means of a two-angle approach is described above. A first approximation of the planar segments directly results from the ray casting algorithm. This approximation can be refined by applying the idea of the polygon inflation method within this plane. Then only two of the three coordinates x_1, x_2 and x_3 are free variables. This procedure is similar to the slicing method from [14]; however the computed plane is usually not orthogonal to one of the coordinate axes.

5.3. Ray casting combined with soft constraints

As demonstrated in [5, 4, 41, 31] and in further publications, the AFS can significantly be reduced by applying soft constraints. A straightforward implementation of soft constraints like unimodality, closure or monotonicity may be inconsistent with the gap-free intersection property which is of central relevance for the ray casting algorithm. Equality constraints can easily be implemented, see Section 4.8 and Remark 4.1.

The key problem with the unimodality and monotonicity constraints is that it is not evident that Theorem 3.3 still holds for the constraint-restricted set $\mathcal{M}$. At least, the property that $\gamma^* x$ is located on the boundary of $\mathcal{M}^+$, see Thm. 3.3, has to be adapted in an appropriate way if additional constraints are used. A thorough analysis of these questions should be a topic of future work - the current work is devoted to the basic nonnegativity constraint.

6. Numerical results

Next the ray casting algorithm is applied to model problems with $s = 3$ and $s = 4$ components. The results for $s = 3$ are compared with the results gained by

[1] We are very grateful to Annekathrin Jürß, University of Rostock, for providing the idea behind this model problem.

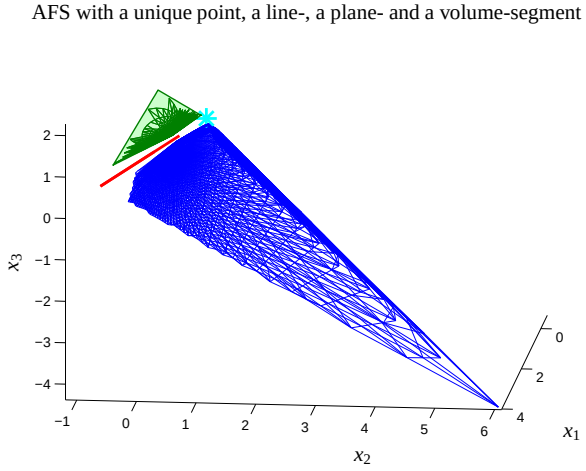

Figure 5: The AFS for the upper triangular rank-4 matrix $D \in \mathbb{R}^{4\times4}$ with $D_{ij} = \max(j - i + 1, 0)$ consists of a unique point (cyan), a linear segment (red), a planar segment (green) and a volume segment (blue).

the polygon inflation method [37, 38] in its *FACPACK* implementation [39].

6.1. Three components

We re-use the model problem of a consecutive irreversible reaction with three components from Sec. 4 of [37]; the unimodal model spectra are strongly overlapping. We use the associated data matrix $D \in \mathbb{R}^{k\times n}$ with $k = 1000$ and $n = 1500$. The AFS consists of three clearly separated segments. The ray casting algorithm is applied with $m = 300$ rays. The bisection method uses the boundary precision $\varepsilon_b = 10^{-3}$, see Eqn. (16). The results are shown in Figure 3 together with the results of the polygon inflation method (by black lines which are for the most part covered by the colored lines). The computation time of the ray casting algorithm is 66.9 seconds. The ray casting algorithm is written in C and the program code uses one core of a 2.4GHz Intel CPU on a standard PC with 16 GB RAM. For this two-dimensional AFS the adaptive polygon inflation method needs only 12 seconds; see also Table 2 in [37].

Next the precision of the results by ray casting is compared to the precision of the polygon inflation method. In order to measure the distance of the two AFS approximations we use the Hausdorff distance $\delta(A, B)$, which is the mutual deviation of two sets A and B

$$\delta(A, B) = \max\left\{\max_{a\in A} d(a, B),\ \max_{b\in B} d(A, b)\right\}. \qquad (21)$$

Therein $d(a, B) = \min_{b\in B} \|a - b\|_2$ is the minimal distance of a point a to a set B. We have computed the three Hausdorff distances separately for each of the three segments of the AFS. The results are

$$(\delta_1,\ \delta_2, \delta_3) = (9.17 \cdot 10^{-3},\ 5.00 \cdot 10^{-3},\ 1.48 \cdot 10^{-2}).$$

For δ_1 (blue segment in Figure 3) and δ_2 (green segment) the distances are consistent with the boundary precision $\varepsilon_b = 10^{-3}$. For δ_3 (red segment in Figure 3) the larger distance can be explained by the efficient spatial adaptivity of the polygon inflation method. In contrast to this, ray casting works with a fixed angular resolution, which can result in a certain corner cut-off, cf. Section 4.3.

6.2. Four components with four isolated segments

Next a model problem with $s = 4$ components, $k = 70$ spectra and $n = 50$ channels is discussed. The factors C and A are shown in the left and centered plot of Figure 6. Ray casting uses $m = 5000$ rays. The computation takes 447.4 seconds. The AFS consists of four isolated segments. Two segments are small and their distance to the origin is larger compared to the two remaining segments. Hence fewer rays hit these smaller segments. For a proper approximation of these two segments we have increased the ray density by the factor 4. The resulting AFS is plotted in Figure 7. The associated series of spectra, the so-called feasible bands, are plotted in Figure 8.

6.3. Four components with a single-segment AFS

By adding constant offsets to the four spectra from Section 6.2, the overlap of these spectra is increased.

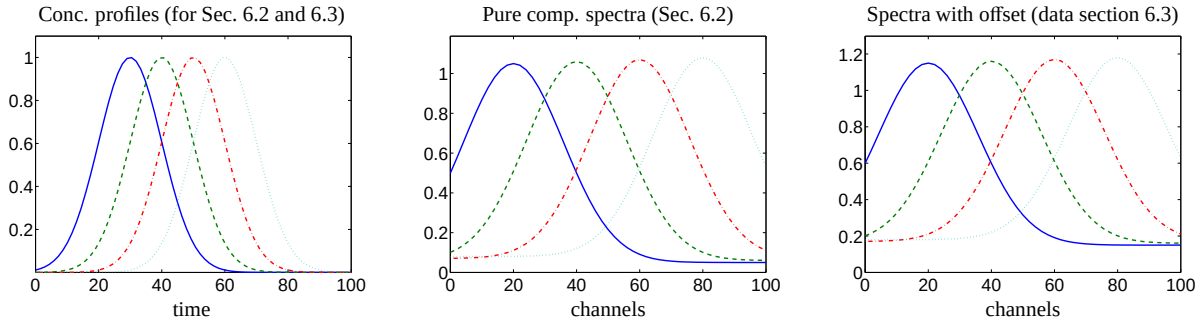

Figure 6: The pure components of the two model problems discussed in Section 6.2 and 6.3. Left: The same concentrational factor C is used for the two model problems. Center: Spectral factor A for the model problem from Section 6.2. The resulting AFS is shown in Figure 7. Right: Spectral factor which underlies Section 6.3 with added offsets which increase the overlap between the spectra. The resulting single-segment AFS is shown in Figure 9.

The factor C remains unchanged; see left subplot in Figure 6. The resulting AFS of $D = CA$ is plotted in Figure 9. The AFS is a topologically connected set with three holes. Compared to the first 3D AFS much more rays hit the AFS. Hence the computation time increases to 1889 seconds.

7. Application to FT-IR experimental data

Next the ray-casting method is applied to experimental and noisy in-situ FT-IR data of a three-component system. The results are compared with those gained by the polygon inflation method.

7.1. Hydroformylation data

The in-situ FT-IR data have been taken from the hydroformylation of 3,3-dimethyl-1-butene with a phosphite-modified rhodium catalyst, see [20]. Here we consider only the spectral window $[1963, 2116]\text{cm}^{-1}$ since this interval contains the characteristic signals from three components, namely the olefin, if the acyl complex and the hydrido complex, see [20]. The wavenumber grid has $n = 639$ channels. With $k = 850$ spectra the matrix D has the dimensions 850×639.

7.2. AFS computations

For this three-component system we compute the AFS for the spectral factor A by means of the ray casting algorithm. The results are compared to those of the polygon inflation algorithm. The ray casting method is run six times with $m \in \{100, 200, 300, 400, 500, 1000, 5000\}$. The boundary precision is $\varepsilon = 10^{-3}$ for all computations and four sweeps of refinement have been applied. The *FACPACK*-implementation of the polygon inflation method is used with the default control parameters $\varepsilon_b = \delta = 10^{-3}$, see [37, 38] for details on the control parameters.

For these computations ray casting and polygon inflation use the same control parameter $\varepsilon = 0.005$ as a lower bound for the acceptable relative nonnegativeness of the factors in the sense

$$\frac{\min\limits_{j=1,\ldots,k} C(j,i)}{\max\limits_{j=1,\ldots,k} C(j,i)} \geq -\varepsilon, \quad \frac{\min\limits_{j=1,\ldots,n} A(i,j)}{\max\limits_{j=1,\ldots,n} A(i,j)} \geq -\varepsilon, \qquad (22)$$

for $i = 1, \ldots, s$.

Figure 10 shows the results of ray casting with $m = 1000$ rays together with the results of polygon inflation. The AFS consists of three isolated segments. The results are almost the same. The Hausdorff distance measure by Eq. (21), which describes the distance between two sets, is applied to the corresponding pairs of AFS segments. The computation times for ray casting for different numbers of rays are listed in Table 1. The polygon inflation method needs only 6.45s. All computations have been run on a standard PC with a 2.4GHz Intel CPU (only one core is used) and 16 GB RAM.

The computation times increase linearly in the number of rays. The Hausdorff distances are consistent with the control parameters on the boundary precision $\varepsilon_b = 10^{-3}$.

These results support that the new ray casting algorithm is a proper tool for AFS computations. However, for two- and three-component systems there are faster methods (e.g. polygon inflation and Borgen plots [7, 32, 19]). The true benefit of ray casting is its easy applicability to systems with four or even more components.

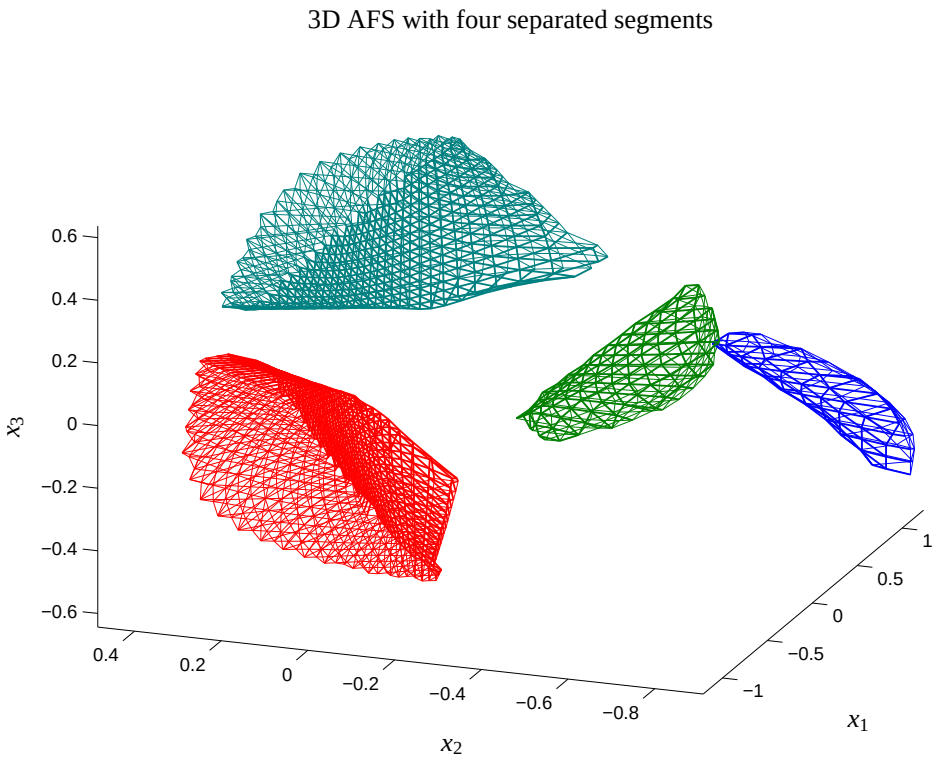

Figure 7: The three-dimensional AFS for the four-component model problem from Section 6.2. The pure component factors C and A are shown in the left and centered plot of Figure 6. The AFS consists of four isolated segments. Only for the two rightmost segments the ray density has been increased by the factor 4 as these segments are more distant from the origin. By means of these additional rays, a comparable quality of surface approximation can be gained for each of the four segments.

8. Conclusion

Up to now various methods have been developed for the geometric construction or numerical approximation of the AFS for two-, three- and four-component systems. Sometimes the geometric construction of the AFS for two- and three-component systems in the form of Borgen plots is considered as the most elegant and approximation-free approach. However, Borgen plots can only be constructed for noise-free and non-perturbed data. Nevertheless, the recent concept of Generalized Borgen plots provides a remedy [19].

For experimentally determined spectral data there is no way around numerical approximation methods. Typically, these numerical methods are tailored to problems for a fixed number of components. For example, the triangle-boundary-enclosure method [12] and polygon inflation [37] work for three-component systems and the slicing approach [12], which extends the triangle-boundary-enclosure method to four-component systems, works for four-component systems. Sometimes the brute-force and computationally most expensive grid search method is considered as a method of first choice since the method is free of any requirements. In some sense the ray casting method can be considered as a smart variant of the grid search method in a sense that grid points are substituted by rays. A decisive benefit of ray casting is that for AFS in d dimensions (for a system with $s = d + 1$ chemical components) a number of m^d grid points must be analyzed but only m^{d-1} must be checked. This saves one dimension and makes ray-casting much faster than the classical grid search.

Our ray casting algorithm is stable for perturbed data. The algorithm can be applied to compute any type of an AFS, e.g. those with isolated segments, those with partially connected segments or those with holes through its outer surface. However, the price for this generality and robustness is a smaller precision-to-effort ratio compared to polygon inflation.

References

[1] H. Abdollahi, M. Maeder, and R. Tauler. Calculation and Meaning of Feasible Band Boundaries in Multivariate Curve Resolu-

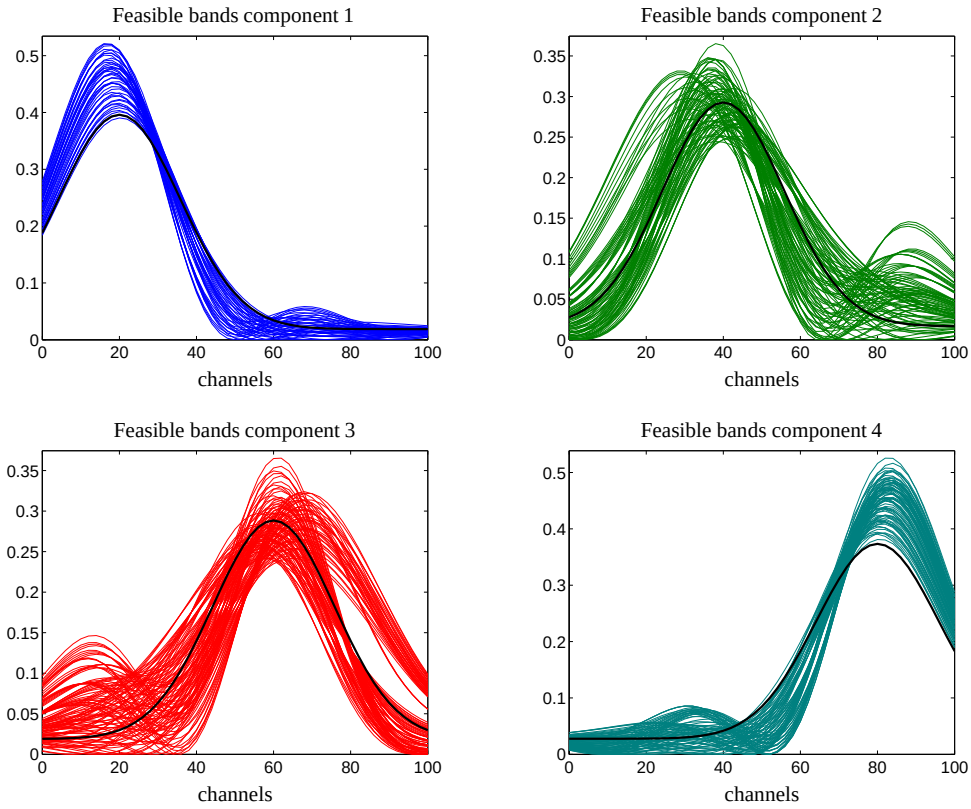

Figure 8: The figure shows the series of spectra, so-called feasible bands, which are associated with the AFS presented in Figure 7. For this four-component system, see Section 6.2, the single spectra are computed along a subset of all computed rays. The distance of points along a certain ray for which spectra are drawn is about 0.1. Spectra are always drawn for the end-points of the intersection of a ray with an AFS segment. The (rescaled) original pure component spectra are drawn by black lines.

tion of a Two-Component System. *Anal. Chem.*, 81(6):2115–2122, 2009.

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

[3] A. Appel. Some techniques for shading machine renderings of solids. In *Proceedings of the April 30–May 2, 1968, Spring Joint Computer Conference*, AFIPS '68 (Spring), pages 37–45, New York, NY, USA, 1968. ACM.

[4] 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.

[5] 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.

[6] L.E. Blumenson. A derivation of n-dimensional spherical coordinates. *The American Mathematical Monthly*, 67:63–66, 1960.

[7] O.S. Borgen and B.R. Kowalski. An extension of the multivariate component-resolution method to three components. *Anal. Chim. Acta*, 174:1–26, 1985.

[8] 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.

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

[10] 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.

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

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

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

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

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

[16] H. Haario and V.M. Taavitsainen. Combining soft and hard modelling in chemical kinetics. *Chemom. Intell. Lab. Syst.*,

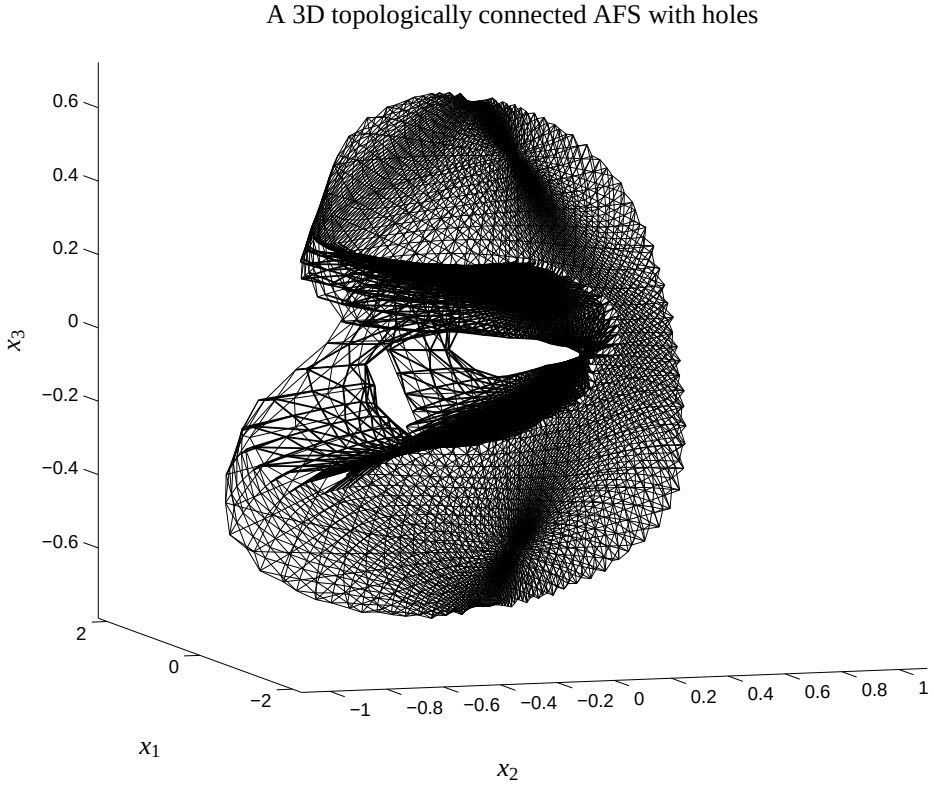

Figure 9: The three-dimensional AFS for the four-component model problem from Section 6.3. By increasing the overlap of the spectra compared to the model problem from Section 6.2, the rotational ambiguity increases together with the volume of the AFS. The AFS is a topologically connected set with three holes through the outer boundary.

44:77–98, 1998.

[17] J.C. Hamilton and P.J. Gemperline. Mixture analysis using factor analysis. II: Self-modeling curve resolution. *J. Chemom.*, 4(1):1–13, 1990.

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

[19] 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.

[20] 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.

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

[22] M. Maeder. Evolving factor analysis for the resolution of overlapping chromatographic peaks. *Anal. Chem.*, 59(3):527–530, 1987.

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

[24] M. Maeder and A. D. Zuberbühler. The resolution of overlapping chromatographic peaks by evolving factor analysis. *Anal. Chim. Acta*, 181(0):287–291, 1986.

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

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

[27] R. Manne. On the resolution problem in hyphenated chromatography. *Chemom. Intell. Lab. Syst.*, 27(1):89–94, 1995.

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

[29] 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.

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

[31] N. Rahimdoust Mojdehi, 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.

[32] 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.

[33] 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.

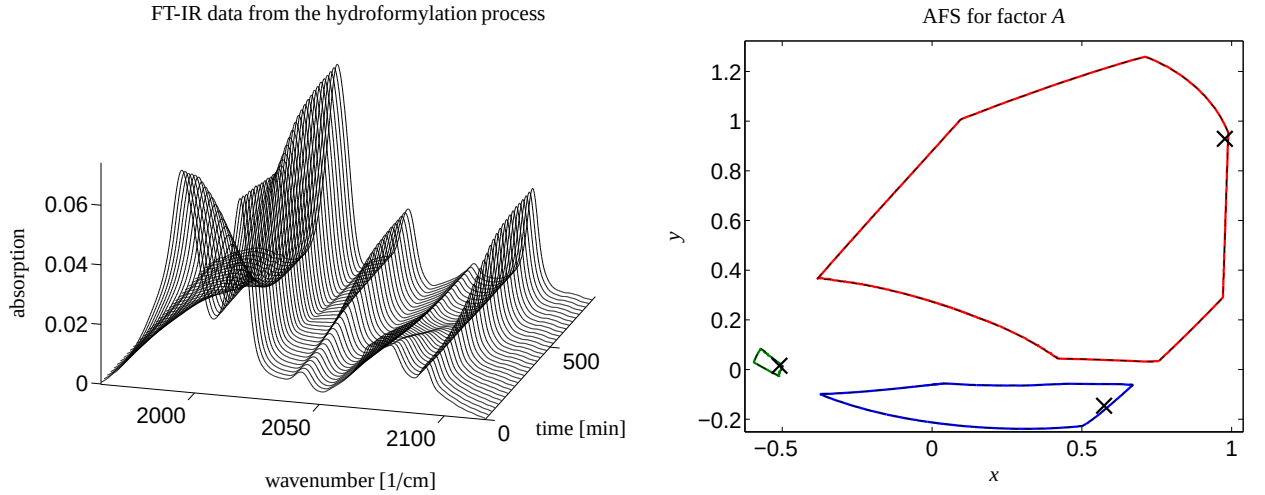

Figure 10: Left: The series of FT-IR data from the hydroformylation process, see Section 7. Right: The AFS computed by the ray casting algorithm with $m = 1000$ rays. The AFS computed by the polygon inflation algorithm is plotted by black dashed lines which are almost covered by the colored lines. The three cross markers represent the true solution which has been computed by means of an additional kinetic model [8, 18, 33, 42].

m	Computation times	δ_1	δ_2	δ_3
100	19.6 s	$1.51 \cdot 10^{-2}$	$3.23 \cdot 10^{-2}$	$4.23 \cdot 10^{-2}$
200	37.1 s	$1.51 \cdot 10^{-2}$	$1.42 \cdot 10^{-2}$	$1.87 \cdot 10^{-2}$
300	57.3 s	$3.84 \cdot 10^{-3}$	$8.20 \cdot 10^{-3}$	$1.08 \cdot 10^{-2}$
400	75.1 s	$5.26 \cdot 10^{-3}$	$5.14 \cdot 10^{-3}$	$6.82 \cdot 10^{-3}$
500	95.3 s	$3.84 \cdot 10^{-3}$	$5.50 \cdot 10^{-3}$	$4.68 \cdot 10^{-3}$
1000	179.6 s	$3.84 \cdot 10^{-3}$	$4.50 \cdot 10^{-3}$	$4.68 \cdot 10^{-3}$

Table 1: Comparison of AFS computations by the ray casting algorithm and the polygon inflation method for $D \in \mathbb{R}^{850\times639}$, see Section 7.1. The ray casting algorithm is used for six levels of resolution. The computation times and the Hausdorff distances of the two AFS approximations are listed. The three numbers $(\delta_1, \delta_2, \delta_3)$ are the three distances for the three AFS segments. Remark: Some of the distances remain unchanged if m is doubled. In such cases, the new rays cannot improve the AFS approximation by ray casting so that the Hausdorff distance to the result by polygon inflation decreases.

[34] 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.

[35] 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 in Resolving Spectral Mixtures, Ed. C. Ruckebusch, chapter ..., page ... Elsevier, Cambridge, 2016.

[36] 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.

[37] 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.

[38] 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.

[39] 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.

[40] 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.

[41] 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.

[42] 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.

[43] 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.

[44] 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.