

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

Mathias Sawall^a, Klaus Neymeyr^{a,b}

^aUniversität Rostock, Institut für Mathematik, Ulmenstraße 69, 18057 Rostock, Germany

^bLeibniz-Institut für Katalyse, Albert-Einstein-Straße 29a, 18059 Rostock, Germany

Abstract

The area of feasible solutions (AFS) of a multivariate curve resolution method is the continuum of feasible solutions under the given constraints. In the current paper the AFS is computed only on the condition of nonnegative solutions. This work is a continuation of a paper (J. Chemometrics 28:106-116, 2013) on the polygon inflation algorithm for AFS computations. In this second part various properties of the AFS are analyzed. First, its boundedness is proved, which is a necessary condition for its numerical computation. Second, it is shown that the origin is never contained in the area of feasible solutions. This fact is the basis for the inverse polygon inflation algorithm, which allows to compute specific types of an AFS containing a hole.

The numerical computation of the AFS is a complicated and computationally expensive process. The construction of proper objective functions for the AFS-optimization problem appears to be decisive. The paper contains a comparative analysis of two objective functions and describes the ideas of the new *FAC-PACK* toolbox for MATLAB. This freely available toolbox contains a numerical implementation of the polygon inflation and of the inverse polygon inflation algorithm.

Key words: factor analysis, pure component decomposition, nonnegative matrix factorization, area of feasible solutions, polygon inflation.

1. Introduction

The area of feasible solutions (AFS) represents the continuum of all solutions of multivariate curve resolution techniques under pre-given constraints. In this paper we consider the AFS only for nonnegativity constraints on the spectral factor and on the concentration factor. The AFS allows to gain an overview on the possible solutions from which an MCR method extracts one final solution by using soft and hard models. However, the reliability of the solution depends on the correctness of the models. A stable and precise numerical computation of the AFS is a time-consuming process, which is made more difficult by noisy data.

The computation of the AFS for a two-component system goes back to Lawton and Sylvestre [1] in 1971. Borgen and Kowalski [2] have extended AFS computations to three-component systems in 1985. For further important contributions see [3, 4, 5, 6] and the refer-

ences therein. We use the term *area of feasible solutions* regardless of the dimension or number of components of the system since we understand the term not in the sense of a two-dimensional surface area but more in the sense of a region or territory.

In 2011 Golshan, Abdollahi and Maeder provided a new idea for the numerical approximation of the boundary of the AFS by a chain of equilateral triangles [7]. An alternative solution for such a numerical approximation of the AFS by a sequence of adaptively refined polygons has been suggested in the first part of this paper [8].

To our knowledge some topological properties of the AFS have never been analyzed. For instance, it is not clear that the AFS is (under some mild assumptions) a bounded set. Boundedness, however, is a necessary prerequisite for a successful numerical approximation. Often the AFS consists of three clearly separated subsets. It can also be a single topologically connected set with a hole; see Figure 1 for two typical areas of feasi-

December 11, 2013

ble solutions. These different topologies require adapted computational strategies. All this justifies the following objectives:

1. analyze some theoretical properties of the Borgen and Kowalski approach to the AFS and their impact on properties of the AFS,
2. compare the two objective functions for the triangle-enclosure algorithm and the polygon inflation algorithm,
3. find approximation schemes which allow to approximate the AFS for typical cases in which the AFS consists of three separated segments or only one segment with a hole,
4. present a fast and stable numerical method for the computation of such one-segment AFS,
5. show how the AFS can be reduced if additional information on the factors is available.

The paper is organized as follows. In Section 2 some mathematical properties of the AFS are analyzed. Section 3 is devoted to a comparative analysis of two objective functions which are key tools for the numerical AFS approximation. In Section 4 the new inverse polygon inflation scheme, which allows to compute an AFS with a hole, is introduced. The inverse polygon inflation algorithm is a central part of the new *FAC-PACK* toolbox in *MATLAB* for AFS computations. In Section 5 techniques are presented which allow to reduce an AFS by means of additional information on the factors. All this is accompanied by various numerical examples in sections 3, 4 and 5.

1.1. Data sets

Within this paper the algorithms and new concepts are tested for the three data sets:

1. Rhodium catalyzed hydroformylation, see [9, 8]. A number of $k = 1045$ FT-IR spectra with $n = 664$ spectral channels is used. There are $s = 3$ independent components namely the olefin, the acyl complex and the hydrido complex. The AFS is shown in Figure 1 (left) and in Figure 6.
2. Formation of hafnacyclopentene, see [10, 11, 12]. A number of $k = 500$ UV/Vis spectra with $n = 381$ spectral channels is given. The AFS for this system with $s = 3$ independent components is shown in Figures 1 (right) and 7.
3. ButiPhane ligands and hydrogenation activity, see [13]. A number of $k = 82$ UV/Vis spectra with $n = 1951$ channels is given. The AFS for this system with $s = 2$ independent components is shown in Figure 3.

The collection of spectra for these three reaction systems are shown together with the concentrations profiles and spectra of the pure components at the end of this paper in Figures 8–10. For more details on these problems see [9], [11] and [13].

2. On the AFS

The Borgen and Kowalski [2] approach to the AFS for three-component systems together with further references and explanations has been introduced in the first part [8] of this paper in Section 2.2. Next only those equations are compiled which are essential for this paper.

The starting point is a $k \times n$ spectral data matrix D whose nonnegative factorization CA is desired. The $k \times s$ concentration matrix C contains the concentration profiles of the s components in its columns, and the $s \times n$ spectral factor A contains the associated s pure-component spectra in its rows. Usually, a given D has an infinite number (or continuum) of nonnegative factorizations. The AFS is a low dimensional representation of these solutions. If D has the rank s , then the truncated rank- s singular value decomposition reads $D = \tilde{U}\tilde{\Sigma}\tilde{V}$ with $\tilde{U} \in \mathbb{R}^{k \times s}$, $\tilde{\Sigma} \in \mathbb{R}^{s \times s}$ and $\tilde{V} \in \mathbb{R}^{n \times s}$. This allows to set $C = \tilde{U}\tilde{\Sigma}T^{-1}$ and $A = T\tilde{V}^T$ where T is an $s \times s$ regular matrix and $D = \tilde{U}\tilde{\Sigma}T^{-1}T\tilde{V}^T$. If D contains perturbed spectral data, then the rank of D is usually larger than s and the upper relations hold in an approximate form.

The AFS for an s -component system is a subset of the \mathbb{R}^{s-1} with the form

$$\mathcal{M} = \{t \in \mathbb{R}^{1 \times s-1} : \text{exists invertible } T \in \mathbb{R}^{s \times s}, T(1, \cdot) = (1, t), \tilde{U}\tilde{\Sigma}T^{-1} \geq 0 \text{ and } T\tilde{V}^T \geq 0\}. \quad (1)$$

For a three-component system this simply reads

$$\mathcal{M} = \{(\alpha, \beta) \in \mathbb{R}^2 : \det(T) \neq 0, C, A \geq 0\}$$

with $C = \tilde{U}\tilde{\Sigma}T^{-1}$, $A = T\tilde{V}^T$ and

$$T = \begin{pmatrix} 1 & \alpha & \beta \\ 1 & s_{11} & s_{12} \\ 1 & s_{21} & s_{22} \end{pmatrix}. \quad (2)$$

2.1. Mathematical foundation

The definition of \mathcal{M} in (1) implies that the rows a of a feasible factor A can be presented by linear combinations of the rows of \tilde{V}^T in the form

$$a = (1, t) \cdot \tilde{V}^T. \quad (3)$$

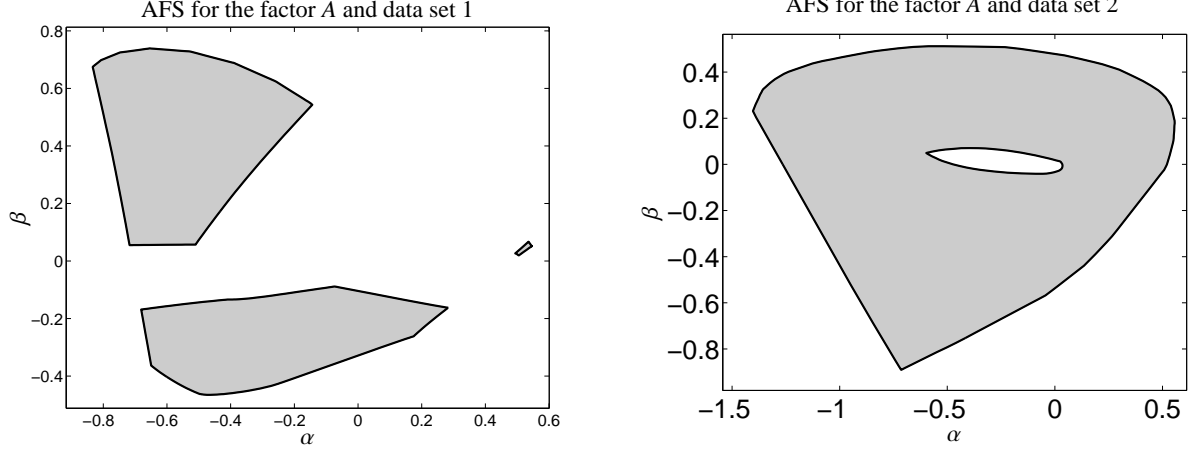


Figure 1: Typical areas of feasible solutions. Left: Three separated segments forming the AFS (Data set 1: hydroformylation process). Right: The AFS is a connected set with a hole (Data set 2: formation of hafnacyclopentene).

The fixed 1 in the row vector $(1, t)$ guarantees that *any* spectrum has a contribution from the first right singular vector. This property is by no means evident and has to be proved. Theorem 2.2 shows that the AFS representation (1) is valid if and only if $D^T D$ is an irreducible matrix. For such matrices Theorem 2.4 shows that \mathcal{M} is a bounded set. Boundedness is a necessary prerequisite for the numerical techniques in [7, 8] to compute the AFS. To our knowledge no proof on the boundedness has been given so far.

The central results of this section are Theorems 2.2 and 2.4. In these theorems $\text{rank}(D) = s$ is assumed with $s \geq 2$. The results do not necessarily hold for perturbed data and if s is smaller than $\text{rank}(D)$. Some implications of these theorems are summarized at the end of this section.

Definition 2.1. Let P be an $n \times n$ permutation matrix, i.e. P is a column permutation of the identity matrix.

An $n \times n$ matrix H with $n \geq 2$ is called *reducible*, if a permutation matrix P exists so that

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

Therein $H_{1,1}$ is an $m \times m$ submatrix and $H_{1,2}$ is an $m \times (n-m)$ submatrix with $1 \leq m < n$. If such a permutation matrix P does not exist, then H is called an *irreducible matrix*.

The next theorem proves that the Borgen and Kowalski approach (with 1s in the first column of T) is justified. Further, the result is used in Theorem 2.4 on the boundedness of the AFS.

Theorem 2.2. Let $D \in \mathbb{R}^{k \times n}$ be a nonnegative matrix with $\text{rank}(D) = s$ which has no zero column. Further let $U\Sigma V^T$ be a singular value decomposition of D and let \tilde{V} be the submatrix of V formed by its first s columns.

There exists no row vector $t \in \mathbb{R}^{1 \times s-1} \setminus \{0\}$ with

$$(0, t) \cdot \tilde{V}^T \geq 0 \quad (4)$$

(in words: any linear combination of the columns 2, ..., s of \tilde{V} has negative components) if and only if $D^T D$ is an irreducible matrix.

Proof. First let $D^T D$ be irreducible. The Perron-Frobenius-theorem [14] guarantees that the first right singular vector $V(:, 1)$ is a sign-constant vector without zero components. This means that either $V(:, 1) > 0$ or $V(:, 1) < 0$. If $V(:, 1) > 0$ and assuming a vector $t \in \mathbb{R}^{s-1} \setminus \{0\}$ satisfying (4), then it holds that

$$\underbrace{(0, t) \cdot \tilde{V}^T}_{\geq 0 \text{ and } \neq 0} \cdot \underbrace{\tilde{V}(:, 1)}_{> 0} > 0.$$

This is a contradiction to the orthogonality of \tilde{V} since its first column $\tilde{V}(:, 1)$ is orthogonal to all the remaining columns $\tilde{V}(:, 2 : s)$. For the case $V(:, 1) < 0$ the arguments are almost the same.

In order to prove the opposite direction by contraposition, let $D^T D$ be a reducible matrix. According to Definition 2.1 there is a permutation matrix P so that

$$PD^T DP^T = \begin{pmatrix} D_1 & 0 \\ 0 & D_2 \end{pmatrix}.$$

The right upper block is also a zero block since $PD^T DP^T$ is a symmetric matrix.

By assumption D has no zero-columns so that all diagonal elements of $D^T D$ are nonzero; i.e., $d_k^T d_k \neq 0$ where d_k is the k -th column of D . Therefore $P D^T D P^T$ has no zero-columns so that D_1 and D_2 are nonzero and nonnegative matrices. Without loss of generality D_1 and D_2 can be assumed to be irreducible matrices; otherwise the argumentation is applied to proper irreducible submatrices.

Let λ_1 and λ_2 be the eigenvalues of D_1 resp. D_2 with the largest modulus. The Perron-Frobenius theorem guarantees that λ_1 and λ_2 are (by irreducibility) simple and positive eigenvalues. The associated normalized eigenvectors u_1 and u_2 are component-wise positive vectors. These eigenvectors are among the right singular vectors of D so that for proper indexes i_1, i_2

$$P\tilde{V}(:, i_1) = \begin{pmatrix} \pm u_1 \\ 0 \end{pmatrix}, \quad P\tilde{V}(:, i_2) = \begin{pmatrix} 0 \\ \pm u_2 \end{pmatrix}. \quad (5)$$

Therein \pm expresses that the orientation of a singular vector is not uniquely determined. Without loss of generality let $i_2 \neq 1$ (otherwise $i_1 \neq 1$). Then let $(0, t) := \pm e_{i_2}^T$ where e_{i_2} is the standard basis column vector whose i_2 -th component equals 1 and all other components are 0. From (5) one gets

$$P\tilde{V}(\pm e_{i_2}) = \begin{pmatrix} 0 \\ u_2 \end{pmatrix} \geq 0.$$

Transposition of this equation and right multiplication by P results in

$$\pm e_{i_2}^T \tilde{V}^T \underbrace{P^T P}_I = (0, u_2^T) P \geq 0.$$

Together with (4) this completes the proof. \square

Corollary 2.3. *Let D^T satisfy the assumptions of Theorem 2.2. Then no $v \in \mathbb{R}^{s-1}$ exists with*

$$\tilde{U}\tilde{\Sigma} \begin{pmatrix} 0 \\ v \end{pmatrix} \geq 0$$

if and only if DD^T is irreducible.

Proof. From $D^T = \tilde{V}\tilde{\Sigma}\tilde{U}^T$ and non-existence of t with $(0, t)\tilde{U}^T \geq 0$ if and only if DD^T is irreducible one gets the non-existence of t with $(0, t)\tilde{\Sigma}\tilde{U}^T \geq 0$. Transposition of the last inequality proves the proposition. \square

The matrix $D^T D$ can be assumed to be irreducible for spectroscopic applications. Otherwise, the series of spectra decomposes into apparently separated or non-coupled subblocks. A trivial example of a reducible matrix is the 3-by-3 identity matrix $D = I_3$ so that

$D^T D = I_3$ for which T is not necessarily in the form (2) (since $V = T = T^{-1} = C = A = I_3 \geq 0$ is a feasible solution).

An important consequence of Theorem 2.2 is that the AFS is a bounded set. The AFS \mathcal{M} is a subset of

$$\mathcal{M}^+ = \{t \in \mathbb{R}^{1 \times s-1} : (1, t)\tilde{V}^T \geq 0\}, \quad (6)$$

which is closely related to FIRPOL in [2, 3]. The set \mathcal{M}^+ stands for the nonnegativity of the spectral factor A only, and \mathcal{M}^+ is the intersection of the n half-spaces

$$\{t \in \mathbb{R}^{1 \times s-1} : t\tilde{V}(i, 2 : s)^T \geq -\tilde{V}(i, 1)\}, \quad i = 1, \dots, n. \quad (7)$$

The next theorem shows that \mathcal{M} and \mathcal{M}^+ are bounded sets for irreducible $D^T D$.

Theorem 2.4. *Let D satisfy the assumptions of Theorem 2.2. Then \mathcal{M}^+ by (6) and \mathcal{M} are bounded if and only if $D^T D$ is an irreducible matrix.*

Proof. Since \mathcal{M}^+ is an intersection of the half-spaces (7), boundedness of \mathcal{M}^+ means that there is no $t \in \mathbb{R}^{1 \times s-1}$ so that according to (6)

$$(1, \gamma t)\tilde{V}^T \geq 0 \quad (8)$$

for all $\gamma \geq 0$ (otherwise \mathcal{M}^+ would be unbounded in the direction t).

Inequality (8) can be rewritten as

$$\gamma t\tilde{V}(:, 2 : s)^T \geq 0 \geq -\tilde{V}(:, 1).$$

The null vector has been inserted in this chain of inequalities which is justified because $\tilde{V}(:, 1)$ is a nonnegative vector. Hence boundedness means that there is no $t \in \mathbb{R}^{1 \times s-1}$ so that $t\tilde{V}(:, 2 : s)^T \geq 0$. Equivalently, there is no t so that $(0, t)\tilde{V} \geq 0$ and thus $D^T D$ is irreducible due to Theorem 2.2. This proves the assertion for \mathcal{M}^+ . As \mathcal{M} is a subset of \mathcal{M}^+ the proof is completed. \square

With few additional assumptions one can show that the set \mathcal{M} does not include the origin (i.e. the zero vector).

Theorem 2.5. *Let $D \in \mathbb{R}^{k \times s}$ be a nonnegative rank- s matrix so that $D^T D$ and DD^T are irreducible matrices and that a factorization $D = CA$ with nonnegative factors exists. Then $0 \notin \mathcal{M}$.*

Further, the first left singular vector $\tilde{U}(:, 1)$ is not the concentration profile of a pure component and the first right $\tilde{V}(:, 1)$ is not the spectrum of one of the pure components.

Proof. Let $D = CA$ be a factorization with $0 \leq C \in \mathbb{R}^{k \times s}$ and $0 \leq A \in \mathbb{R}^{s \times n}$. According to (1) the first row of A equals $A(1, :) = (1, t)\tilde{V}^T$ for some $t \in \mathcal{M}$. The concentration profile of the second component is $C(:, 2) = \tilde{U}\tilde{\Sigma}v$ for $v \in \mathbb{R}^s$. Corollary 2.3 proves that $v_1 \neq 0$.

Thus $(1, t)$ is the first row of T and v is the second column of T^{-1} . From

$$0 = I_{1,2} = (TT^{-1})_{1,2} = (1, t) \cdot v$$

one derives $0 \neq v_1 = -t \cdot v(2 : s, 1)$. Thus neither t nor $v(2 : s, 1)$ are null vectors. So $0 \notin \mathcal{M}$ and $\tilde{V}(:, 1)$ is not equal to a pure component spectrum. Applying the arguments to D^T shows that $\tilde{U}(:, 1)$ is not a concentration profile of a pure component. \square

3. Objective functions

The numerical computation of the AFS for $s = 3$ and even larger s is a complicated and computationally expensive process. There is no closed-form-representation of the AFS which could simply be drawn by the evaluation of a function. Noisy data makes the computation of the AFS even more difficult.

For three-component systems Borgen and Kowalski [2] as well as Rajkó and István [3] presented for noisy-free data a geometric approach to the construction of the AFS. A direct numerical approximation of the AFS for three-component systems and noisy data can be computed by means of the triangle-enclosing algorithm [7] and the polygon inflation algorithm [8]. These iterations aim at an approximate computation of the boundary of the AFS. The central process behind these algorithms is a routine which decides whether a certain point (α, β) is contained in \mathcal{M} . Such points are called *valid*. Points exterior to \mathcal{M} are *non-valid* points.

The algorithms in [7] and [8] make use of different objective functions. Next these objective functions are compared for general s -component systems. The first objective function requires a more expensive computational process compared to the second function. The two algorithms give the same results for nonnegative data. In case of noisy data the resulting AFS approximations may slightly differ.

3.1. The AFS for s -component systems

According to (1) and (2) the problem for an s -component system is to find regular matrices

$$T = \begin{pmatrix} 1 & t_1 & \dots & t_{s-1} \\ 1 & & & \\ \vdots & & \mathbf{S} & \\ 1 & & & \end{pmatrix} \in \mathbb{R}^{s \times s} \quad (9)$$

with a row vector $t = (t_1, \dots, t_{s-1}) \in \mathbb{R}^{1 \times s-1}$ and a submatrix $S \in \mathbb{R}^{(s-1) \times (s-1)}$ in a way that $C = \tilde{U}\tilde{\Sigma}T^{-1}$ and $A = T\tilde{V}^T$ are nonnegative matrices. In order to decide whether t is valid ($t \in \mathcal{M}$) or non-valid ($t \notin \mathcal{M}$) one has to solve an optimization problem. The solution is either a proper submatrix S or its non-existence can be stated. We call this process the *AFS-decision* over (t, S) .

3.2. The ssq approach

The AFS-decision over (t, S) in [4, 5, 7] is based on

$$\text{ssq} : \mathbb{R}^{s-1} \times \mathbb{R}^{(s-1) \times (s-1)} \rightarrow \mathbb{R}, \quad (t, S) \mapsto \|D - C_+A_+\|_F^2$$

with

$$C_+ = \max(\tilde{U}\tilde{\Sigma}T^{-1}, 0), \quad A_+ = \max(T\tilde{V}^T, 0)$$

and T given by (9). Further $\|\cdot\|_F$ denotes the Frobenius norm [15].

By using ssq the AFS \mathcal{M} is

$$\mathcal{M} = \left\{ t \in \mathcal{M}^+ : \min_{S \in \mathbb{R}^{(s-1) \times (s-1)}} \text{ssq}(t, S) \leq \epsilon \right\} \quad (10)$$

with $\epsilon = 0$. For practical computations in presence of rounding errors one can use $\epsilon = 10^{-12}$. The minimization problem in [7] has been solved by the Nelder-Mead simplex minimization (MATLAB routine `fminsearch`). This approach is summarized as follows:

Objective function 1. *The AFS-decision over (t, S) results in a valid t , i.e. $t \in \mathcal{M}$, if the minimization of the objective function ssq yields 0 up to rounding errors.*

3.3. Alternative approach to the AFS-decision

Next an alternative minimization problem is discussed with a much smaller number of squares. A first rapid test on the validity of $t \in \mathcal{M}$ can result in an additional acceleration of the computational process.

We start with this rapid test for a given t . Nonnegativity of $T\tilde{V}^T$ implies that

$$-t \cdot V(i, 2 : s)^T \leq V(i, 1), \quad i = 1, \dots, n. \quad (11)$$

This test does not require the computation of S in (9) so that the computational costs for checking (11) are negligible.

If a certain row vector t has passed the test (11), then we consider the function

$$f : \mathbb{R}^{s-1} \times \mathbb{R}^{(s-1) \times (s-1)} \rightarrow \mathbb{R}^{ks+n(s-1)+1}, \quad (12)$$

$$(t, S) \mapsto \begin{pmatrix} \min(0, C_{il}) \\ \min(0, A_{lj}) \\ \|I_s - T^+T\|_F^2 \end{pmatrix} \begin{matrix} ks \text{ components} \\ n(s-1) \text{ components} \\ 1 \text{ component.} \end{matrix}$$

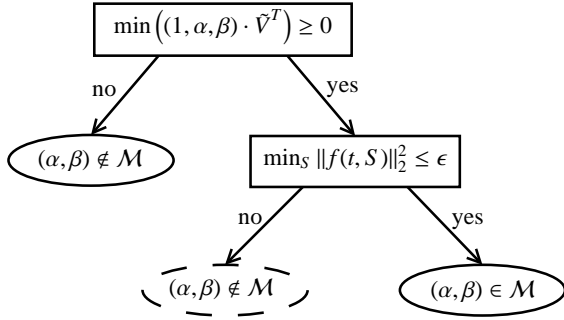


Figure 2: Decision tree for objective function 2. The dashed ellipse expresses some uncertainty as the numerical minimization concerning S may fail, cf. Section 3.5.

The first ks components of f are either equal to 0 or the negative components of $C \in \mathbb{R}^{k \times s}$. The following $n(s-1)$ components of f are either equal to 0 or the negative components of $A \in \mathbb{R}^{s \times n}$; the first row of A has not to be checked on negative components due to the test (11). The last component of f aims at avoiding a rank-deficient matrix T with $\text{rank}(T) < s$.

The least-squares minimization of f includes a number of $ks + n(s-1) + 1$ squares. Since s is a fixed small number and k as well as n are potentially large the total costs for minimizing (12) increases as $O(k+n)$ wherein O is the Landau symbol. In contrast to this, the ssq approach is more expensive with $O(kn)$. For the numerical minimization of $\|f(t, S)\|_2^2$ a code for nonlinear least-squares minimization can be applied like `lsqnonlin` in `MATLAB`. For our *FAC-PACK* implementation we use the FORTRAN code `NL2SOL` [16].

With (12) this allows to define the area of feasible solutions as follows

$$\mathcal{M} = \left\{ t \in \mathcal{M}^+ : \min_{S \in \mathbb{R}^{(s-1) \times (s-1)}} \|f(t, S)\|_2^2 \leq \epsilon \right\}.$$

Theoretically $\epsilon = 0$ but we use $\epsilon = 10^{-10}$ for practical computations.

Objective function 2. *The AFS-decision over (t, S) results in a valid t , i.e. $t \in \mathcal{M}$, if the test (11) is passed and the minimization of the objective function f results in 0 up to rounding errors. The decision tree is shown in Fig. 2.*

3.4. Negative components

Noisy spectral data or negative components of the spectral data matrix D due to some data preprocessing does not always allow to find nonnegative matrix factors

C and A . A proper treatment of small negative components of these matrices appears to be crucial for a successful construction of the AFS. The approaches from Sections 3.2 and 3.3 treat such negative components differently.

Let us first discuss the limit case of $\min_{S \in \mathbb{R}^{(s-1) \times (s-1)}} f(t, S) = f(t, S^*) = 0$. Then by definition of f , see Equation (12), it is guaranteed that C and A are nonnegative matrices. Thus

$$\max(C, 0) \cdot \max(A, 0) = C \cdot A = D.$$

This implies that $\text{ssq}(t, S^*) = 0$. It is not clear that the other case of a vanishing ssq of $D - C_+ A_+$ implies that f also vanishes. To show this, the reconstruction $D = C_+ A_+$ with the truncated matrices C_+ and A_+ must imply $C_- = 0$ and $A_- = 0$. It is not clear if such properties can be proved. But as C_+ and A_+ are feasible nonnegative factors they are in any case represented by the AFS.

Here we accept the factors C and A if the relative negative portion in the columns of C and the rows of A is bounded as follows

$$\frac{\min_j C_{ji}}{\max_j |C_{ji}|} \geq -\epsilon, \quad \frac{\min_j A_{ij}}{\max_j |A_{ij}|} \geq -\epsilon$$

for $i = 1, \dots, 3$. The parameter ϵ , see Section 3.3, is taken as $\epsilon = 10^{-12}$ for model problems without perturbations.

Similarly the first rapid test (11) is reformulated as

$$\frac{V_{i1} + t_1 V_{i2} + \dots + t_{s-1} V_{is}}{\|(1, t) \cdot V^T\|_\infty} \geq -\epsilon, \quad i = 1, \dots, n.$$

Therein $\|\cdot\|_\infty$ is the maximum norm, i.e. the maximum of the absolute values of the components.

Further, the first ks components of f are taken as

$$\min(0, C_{i\ell} / \|C(:, \ell)\|_\infty + \epsilon)$$

and the following $n(s-1)$ components are similarly

$$\min(0, A_{\ell j} / \|A(\ell, :)\|_\infty + \epsilon).$$

3.5. Reliability of the numerical optimization

The AFS-decision for a vector $t \in \mathbb{R}^{s-1}$ is a numerically expensive and potentially instable process since for the given t a proper S is to be computed by solving an optimization problem. This numerical computation may fail. In particular such problems are to be expected if only a poor initial approximation for S is available at the start of the optimization procedure.

However, for two-component systems the situation is very simple, see Section 3.6. For $s \geq 3$ the decision tree

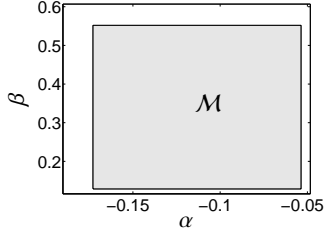


Figure 3: AFS for a two component system for the ButiPhane data.

shown in Figure 2 applies. The rapid test on nonnegativity by inequality (11) appears to be non-critical. Further, any decision $t \in \mathcal{M}$ can be trusted as the optimization process has found a solution. A decision $t \notin \mathcal{M}$ due to $\min_S \|f(t, S)\|_2^2 > \epsilon$ may not be trusted as, e.g., the optimization procedure may have got stuck in a local and non-global minimum. To reduce the risk of such misclassifications we always take special care to generate good initial values for S . Such initial matrices can be computed by local averaging final and trusted matrices S on points in a close neighborhood of t .

3.6. The AFS for two-component system

For a two-component system the matrix T is simply given by

$$T = \begin{pmatrix} 1 & \alpha \\ 1 & \beta \end{pmatrix}.$$

If we ignore noise and let $\epsilon = 0$, then using the objective function 2 results in

$$(\alpha, \beta) \in [a, b] \times [c, d] \quad \text{or} \quad (\beta, \alpha) \in [a, b] \times [c, d]$$

with

$$\begin{aligned} a &= - \min_{i: V(i,2) > 0} \frac{V_{i1}}{V_{i2}}, & b &= \min_i \frac{U_{i2}\sigma_2}{U_{i1}\sigma_1}, \\ c &= \max_i \frac{U_{i2}\sigma_2}{U_{i1}\sigma_1}, & d &= - \max_{i: V(i,2) < 0} \frac{V_{i1}}{V_{i2}}. \end{aligned}$$

Thus the AFS for a two-component system consists of two real intervals. These two intervals are usually represented by the sides of a rectangle, cf. [5]. A point within this rectangle allows a simultaneous representation of the AFS for the spectral and for the concentration factor (which is fundamentally different from the AFS for three-component systems where either the AFS for the spectral factor or the AFS for the concentration factor is represented in 2D). It is also clear that $b < 0 < c$ and that either $\alpha < 0$ and $\beta > 0$ or alternatively $\alpha > 0$ and $\beta < 0$. All this is consistent with $0 \notin \mathcal{M}$ which is

dimension	objective fct. 1		objective fct. 2
k	n	time [s]	time [s]
75	50	9.15	1.23
150	100	33.31	1.68
300	200	133.04	2.82
450	300	271.95	3.47
750	500	803.86	5.45
1500	1000	3773.88	11.26

Table 1: Total computing times for programs using the two objective functions for the model from [8]. In each of these cases the AFS \mathcal{M}_C and the AFS \mathcal{M}_A are computed.

proved in Theorem 2.5. In case of noisy data, the AFS is still a rectangle. Then the scalars a, b, c and d can be computed numerically by using the bisection method. For instance a and d are the minimal or maximal value of α so that

$$\frac{(1, \alpha)V^T}{\|(1, \alpha)V^T\|_\infty} \geq -\epsilon$$

and b, c are computed analogously by using linear combinations of $U\Sigma$. Figure 3 shows the two-dimensional AFS for the ButiPhane data set, which is data set 3 in Section 1.1.

3.7. Computational costs

Next a direct comparison of the computational costs is given for the two objective functions presented above. First we consider a three-component model problem and second we use FT-IR spectral data from the hydroformylation process. The two objective functions are each applied within the polygon inflation algorithm in order to present in detail the effect of the choice of the objective functions.

3.7.1. A model problem

The three-component model problem is taken from part I of this paper; see Section 4.1 in [8]. Here we consider a series of different values for k , the number of spectra, and n , the number of spectral channels, see Table 1. The last two columns of Table 1 give the total computational times for programs using the objective functions 1 and 2. For these computations the parameters $\epsilon = 10^{-12}$ and $\epsilon_b = \delta = 10^{-3}$ have been used; see Part I for the explanation of ϵ_b and δ .

Figure 4 shows the computation time against kn in a log-log plot. The computational costs for the first objective function increases with $O(kn)$. The second objective function results in a much faster method. The numerical data are consistent with costs increasing as $O(k+n)$.

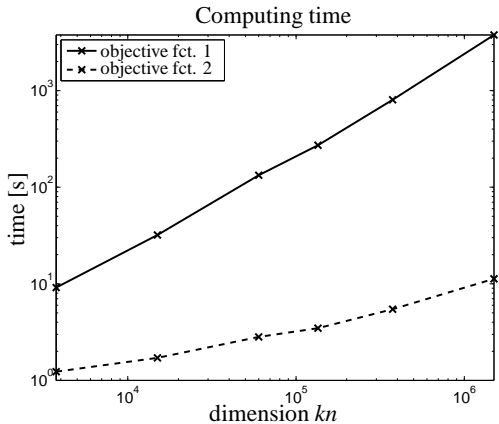


Figure 4: Bilogarithmic plot of the data from Table 1.

AFS	objective fct. 1 time [s]	objective fct. 2 time [s]
\mathcal{M}_C	998.32	4.58
\mathcal{M}_A	862.33	4.39

Table 2: Computing times for the two objective functions for the Rhodium catalyzed hydroformylation. The times for \mathcal{M}_C and \mathcal{M}_A are tabulated separately.

3.7.2. Rhodium catalyzed hydroformylation

Next we consider the Rhodium catalyzed hydroformylation as given by the data set 1 in Section 1.1. Now the parameters are $\varepsilon = -0.01$ and $\varepsilon_b = \delta = 10^{-3}$. The computational times in Table 2 confirm that the second objective function results in a much faster method. The resulting areas of feasible solutions are almost identical if the same value ε_b for the boundary precision is used. The associated Hausdorff distances for the three segments of the AFS are given in Table 1 of [8].

4. Inverse polygon inflation for an AFS with a hole

As shown in Section 2 the AFS is a bounded set which does not include the origin. A challenging question is: *What is the number of isolated segments an AFS may consist of?*

For $s = 2$ -component systems the AFS consists of $p = 2$ separated intervals which are taken as the sides of a rectangular for its presentation; see [1, 5] or Section 3.6. For $s = 3$ -component systems experimental data and model data show that a number of $p = 1$, $p = 3$ or even $p = 6$ segments may occur. We plan to give a mathematical proof that an AFS with two segments cannot occur in a forthcoming publication. A further

interesting question is to show that the number of AFS segments of \mathcal{M}_A and \mathcal{M}_C coincide.

It is well known that the AFS for three-component systems may consist of only one segment and that this segment can contain a hole, which surrounds the origin. Next we describe a variation of the polygon inflation algorithm which can be used to compute such AFS with a hole or an AFS with more than three isolated segments.

4.1. Inverse polygon inflation

If for a three-component systems ($s = 3$) the entire AFS consists of one segment with a hole, then the triangle-enclosure algorithm [7] and the polygon inflation algorithm are to be modified properly. For the triangle-enclosure algorithm two runs are necessary in order to cover the interior and the exterior boundary curve by sequences of triangles.

For the polygon inflation algorithm the exterior polygon and the interior polygon are to be treated differently but the geometric concept of inflation polygon is in each of these cases the same. Only the objective functions are changed.

First, the computational effort to compute the exterior polygon is very small since only $(1, t)\tilde{V}^T \geq 0$ is to be tested according to Equation (6), cf. FIRPOL in [2]. This exterior polygon is just the boundary of the set \mathcal{M}^+ defined in (6). The remaining conditions on t to be a valid vector, i.e. $t \in \mathcal{M}$, are used to define a further set

$$\mathcal{M}^* = \{t \in \mathbb{R}^{s-1} : \min_S \|f(t, S)\|_2^2 = 0\} \quad (13)$$

whose inner boundary is computed by the standard polygon inflation algorithm. For $t \in \mathcal{M}^*$ the definition of \mathcal{M}^* guarantees that T is regular, $C \geq 0$ and $A(2 : 3, :) \geq 0$. The intersection of \mathcal{M}^+ and \mathcal{M}^* , which combines the conditions, results in the AFS

$$\mathcal{M} = \mathcal{M}^+ \cap \mathcal{M}^*. \quad (14)$$

In order to avoid any misinterpretation we mention that \mathcal{M}^* is very different to INNPOL as used in [2, 3].

The algorithm to compute a polygon which approximates the boundary of \mathcal{M}^+ uses an objective function which guarantees (7) to hold. The starting point is the origin which is always in \mathcal{M}^+ . The polygon inflation starts with a triangle enclosing the origin and whose vertices are located on the boundary of \mathcal{M}^+ .

After this the interior boundary of \mathcal{M}^* is computed by using the objective function (12). Therefore the complement $\mathbb{R}^2 \setminus \mathcal{M}^*$ is approximated from the interior of this set. The starting point is, once again, the origin since Theorem 2.5 guarantees that $(0, 0) \notin \mathcal{M}^*$. The

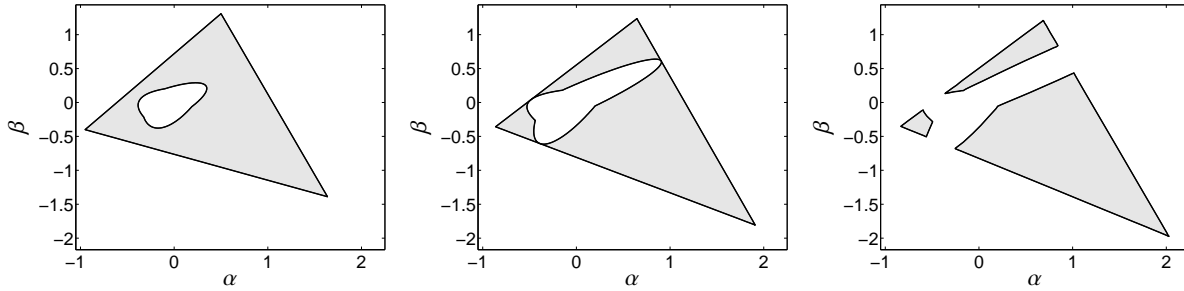


Figure 5: Breaking-up of the AFS \mathcal{M}_A from one to three segments for the Example 4.1 with $\omega_1 = 1.5$. Left: Only one segment with a hole for $\omega_2 = 0.5$. Center: The hole touches each side of the outer triangle for $\omega_2 = 0.26787$. Right: For $\omega_2 = 0.2$ the AFS is broken into three clearly separated segments.

usage of the complement is the reason why we call the algorithm an *inverse polygon inflation*. The computation of the hole of the AFS by applying the polygon inflation to $\mathbb{R}^2 \setminus \mathcal{M}^*$ has the advantage that only few lines of program code are to be adapted. Further, the relevant regions of the two sets \mathcal{M}^+ and \mathcal{M}^* can be computed in a stable way.

If *FAC-PACK* uses the standard polygon inflation algorithm and finds an AFS segment which has a nonzero intersection with at least three quadrants of the Cartesian coordinate system, then the algorithm automatically switches to the inverse polygon inflation. Thus a one-segment AFS is automatically computed by inverse polygon inflation.

4.2. *FAC-PACK* software for AFS computations

The polygon inflation algorithm and the inverse polygon inflation algorithm are implemented in the software toolbox *FAC-PACK* for MATLAB [17]. *FAC-PACK* contains a graphical user interface from which the spectral data can be loaded, a singular value decomposition and an initial nonnegative matrix factorization (NMF) can be computed. This NMF is the starting point for the computation of the AFS. *FAC-PACK* allows to display the spectral and concentration factors which are associated with any points of the AFS. Further, one or even two components can be locked and the resulting reduced AFS is shown.

The *FAC-PACK* homepage containing the software and a tutorial can be accessed at

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

4.3. Weakly separated segments of the AFS

If the segments of an AFS are only weakly separated (in a sense that the polygon inflation algorithm tends to glue separated segments of the AFS to a joint segment; see the AFS in the mid of Figure 5), then the numerical

computation of \mathcal{M}^+ and \mathcal{M}^* , which is followed by their intersection, is a stable and favorable way to construct \mathcal{M} . The inverse polygon inflation procedure can even be applied to general situations with well separated segments - a situation we have often found for FT-IR spectral data. However, the computational procedure for the inverse polygon inflation is somewhat more expensive compared to the direct computation of the three separated segments.

4.4. AFS dynamics showing the separation process

In order to study the possible shapes of an AFS we illustrate the transition of the AFS from one topologically connected segment to an AFS formed by three segments for a three-component model problem. In this separation process three segments which pairwise touch in one point split up to three clearly separated segments. The case of only two separated segments is not observed; we will give a mathematical proof on this in a forthcoming paper. The model problem depends on two real parameters ω_1 and ω_2 .

Example 4.1. Let $D \in \mathbb{R}^{3 \times 3}$ be the nonnegative matrix

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \omega_1 \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} + \omega_2 \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Nonnegative factorizations $D = CA$ and the AFS \mathcal{M}_A are to be computed.

For our computations the parameter ω_1 is fixed to $\omega_1 = 1.5$ and ω_2 is used as a variable. Figure 5 shows that the AFS for $\omega_2 = 0.5$ is one topologically connected set. For $\omega_2 = 0.26787$ the AFS consists of three segments which pairwise touch in one point. For $\omega_2 = 0.2$ one gets three clearly separated segments.

Various numerical experiments indicate that the AFS \mathcal{M}_A for three-component systems is either formed by

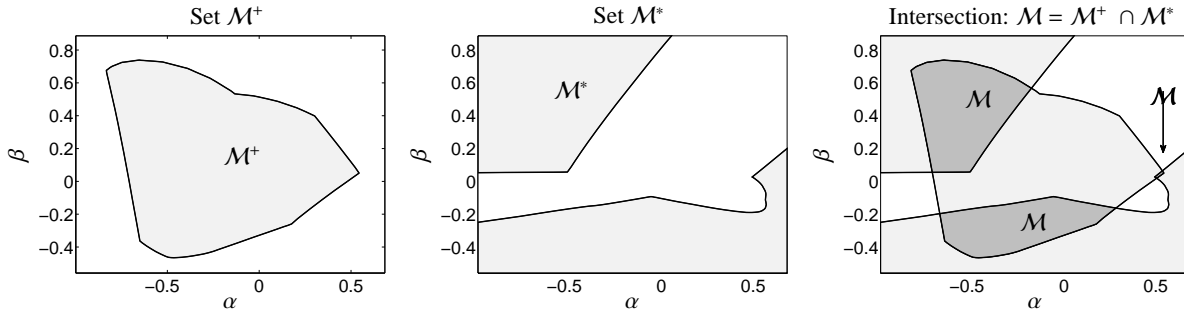


Figure 6: Results of inverse polygon inflation for the hydroformylation data set 1 from Section 1.1. Left: The boundary of \mathcal{M}^+ according to Eq. (6) is similar to FIRPOL [2]. Center: The set \mathcal{M}^* according to (13) is shown in the same α, β -window for which \mathcal{M}^+ is computed. This avoids useless computations since \mathcal{M} is a subset of \mathcal{M}^+ . Right: The AFS \mathcal{M} is the intersection $\mathcal{M} \cap \mathcal{M}^*$.

one topologically connected segment or by three isolated segments. The same holds for the AFS \mathcal{M}_C . Specially constructed matrices D may have much more isolated segments.

4.5. The AFS for spectroscopic data

1. Data set 1 - the Rhodium catalyzed hydroformylation: The reference to this FT-IR spectroscopic data is given in Section 1.1. The spectral AFS for this problem is computed by inverse polygon inflation. Figure 6 shows the sets \mathcal{M}^+ (left), \mathcal{M}^* (center) and their intersection \mathcal{M} , see also Figure 1.
2. Data set 2 - formation of hafnacyclopentene: The reference to this UV/Vis data set is given in Section 1.1. The AFS $\mathcal{M} = \mathcal{M}_A$ is shown in Figure 1 (right plot) and is one topologically connected set with a hole, which contains the origin. This relatively large AFS imposes only weak restrictions on the factor A since the associated spectra are nowhere equal to zero. All this allows a wide range of feasible transformation making the AFS large.

4.6. How to compute line-shaped segments of the AFS

An isolated segment of the AFS is most often either a set whose (mathematical surface) area is larger than zero or it is a single point. In some cases an isolated subset of the AFS appears to be a straight-line segment. In absence of rounding errors and perturbations its surface area equals zero; for slightly perturbed data such a segment practically is a long and narrow band. The polygon inflation method needs some algorithmic enhancement in order to compute such straight-line segments.

In *FAC-PACK* we use an angle-search method for approximating such AFS segments. The starting point is a feasible coordinate $x = (\alpha, \beta)$ as computed by the

NMF. Together with a small radius parameter r the function

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

is considered in order to compute a feasible angle φ so that $g_{r,x}(\varphi) \in \mathcal{M}$. The numerical minimization of the objective function (12) is executed only if the rapid test (11) is passed. The initial x might be one of the endpoints of the line segment or between them. In the latter case and if φ represents a feasible direction, then also $\varphi - \pi$ stands for a feasible direction. For these two oppositely oriented direction maximal values r_l and r_r are computed by the bisection method so that the desired line segment equals the union $\mathcal{L}_l \cup \mathcal{L}_r$ with

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

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

5. How to use additional information on the factors

Nonnegativity of the factors is the only restriction which underlies the construction of the AFS. However, sometimes additional information on the factors is available. This information may consist of the knowledge on certain pure-component spectra or concentration profiles. Alternatively, some isolated peaks within a spectrum may be known. In a recent work Beyramysoltana, Rajkó and Abdollahi [18] have shown a correlation of known factors for three-component systems in the spectral space with lines in the concentration space and vice versa. Comparable results for two-component systems are analyzed by Rajkó [19].

The software toolbox *FAC-PACK* allows to lock certain spectra or concentration profiles and then to compute the reduced and smaller AFS for the remaining components [17]. In this section we focus on the mutual effect of reducing the AFS for A on the AFS for C and vice versa.

Further sources of additional information on the factors might be window factor analysis (WFA) techniques or the evolving factor analysis [20, 21]. In this context the theorems of Manne [22] might help to extract concentration profiles of certain components. The implementation of such local rank information as a part of *FAC-PACK* is a point for the future work. One can also use constraints like unimodality or the proximity of a spectrum in the AFS to a pre-given spectrum in order to reduce the ambiguity. However, in this paper we concentrate on computing the AFS under nonnegativity constraints and we do not want to dilute this approach by introducing too much adscititious information.

5.1. Three-component systems

To discuss a typical and concrete problem we assume that the first spectrum $A(1, :)$ of a three-component system is known. As shown in [12] by coupling and complementarity theorems the knowledge of this spectrum imposes an affine-linear constraint on the associated column $C(:, 1)$ and further linear constraints on the remaining columns $C(:, 2 : s)$. So the knowledge of $A(1, :)$ considerably reduces the original AFS for A . Additionally, the points representing the complementary concentration profiles $C(:, 2 : 3)$ are restricted to a straight line through the AFS for C . The details are as follows: A known spectrum $A(1, :)$ is represented by the coordinates $(t_2/t_1, t_3/t_1)$ in the spectral AFS \mathcal{M}_A with $t = (t_1, t_2, t_3) = A(1, :) \cdot V(:, 1 : s)$. Theorem 4.2 in [12] shows, that the concentration profiles $C(:, 2)$ and $C(:, 3)$ are elements of the subspace

$$\{U\Sigma y : t \cdot y = 0\} = \{U\tilde{y} : t \cdot \Sigma^{-1} \cdot \tilde{y} = 0\}.$$

Together with the required scaling $\tilde{y}_1 = 1$ and using (α, β) as coordinates in \mathcal{M}_C this subspace reads

$$\begin{aligned} L &= \{(\alpha, \beta) : t \cdot \Sigma^{-1} \cdot (1, \alpha, \beta)^T = 0\} \\ &= \left\{ (\alpha, \beta) : \frac{t_2}{\sigma_2} \alpha + \frac{t_3}{\sigma_3} \beta = -\frac{t_1}{\sigma_1} \right\}. \end{aligned}$$

Thus L is a straight line through \mathcal{M}_C and the (α, β) -representatives of the complementary concentration profiles $C(:, 2)$ and $C(:, 3)$ are located on this line.

This reduction process can be continued if a second spectrum $A(3, :)$ is pre-given. For the concentration factor C this means that points in the AFS which represent

the two concentration profiles $C(:, 1 : 2)$ are located on a further straight line. Together with the first straight line the intersection of these lines uniquely determines the second concentration profile $C(:, 2)$. This situation has explicitly been discussed in [12]; see there Theorem 4.2, Corollary 4.3 and Section 6.1. All these restrictions can be combined with any further information/constraints on C and A . However, if $C(:, i)$ and $A(i, :)$ for a certain i are pre-given simultaneously, then the whole problem can be reduced by this component.

5.2. Numerical example

Next we consider UV/Vis spectra for the formation of hafnacyclopentene given by data set 2 in Section 1.1; see also Section 6.1 of [12]. For this three-component system the initial concentrations of the consecutive reaction $X \rightarrow Y \rightarrow Z$ are $c_X(0) = 0.01309$ and $c_Y(0) = c_Z(0) = 0$. The last spectrum of this series is a good approximation of the spectrum of the reaction product Z as the reaction is more or less completed. Together with the first spectrum two pure component spectra are available.

The AFS and its reductions are presented in three steps:

1. The initial areas of feasible solutions for C and A are shown in Figure 7 in the left column. No adscititious information has been used for the AFS computation.
2. If $A(:, 1)$ for the reactant X is given, then the initial AFS can be reduced to two smaller segments, which are shown in Figure 7 in a darker gray. The bold straight line in the AFS for C in Figure 7 (second row and second column) covers the complementary concentration profiles $C(:, 2 : 3)$.
3. If finally the spectrum $A(:, 3)$ for Z is also given, then the resulting AFS is reduced to the right lower AFS segment which is shown in Figure 7 by the darkest gray. A second straight line is added to the AFS for C . Figure 7 (second row and third column) shows this by another bold line. The intersection of the two lines uniquely determines the concentration profile $C(:, 2)$.

5.3. Remark on the AFS for multi-component systems

The AFS for a system with s independent components is a bounded set in the $s - 1$ -dimensional space according to Equation (9). Additional information on the factors can be applied to such a higher-dimensional AFS in a way which is comparable to the techniques explained above.

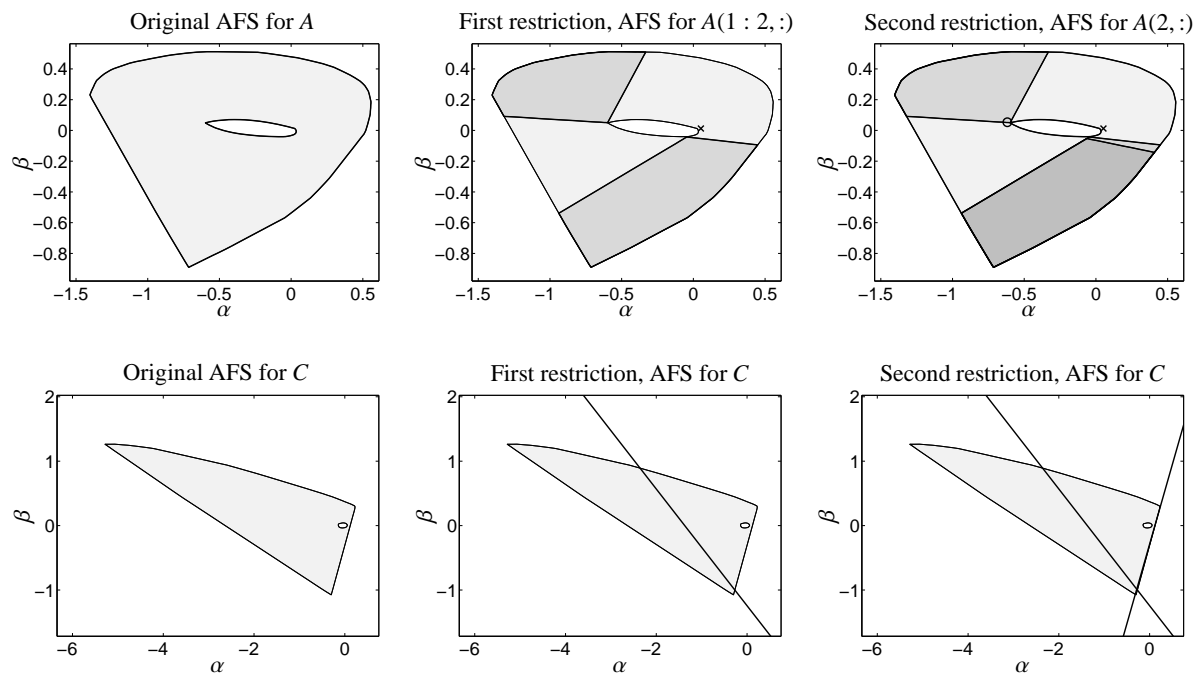


Figure 7: Restriction of the AFS. First row: AFS for factor A. Second row: AFS for factor C. Left column: AFS without any restrictions. Central column: the pure component spectrum of the component Z is marked by \times in the AFS for A. For further explanations see Section 5.2. Right column: the pure component spectrum of the component X is marked by \circ in the AFS for A. For further explanations see Section 5.2.

For example consider a four-component system. If a certain spectrum $A(1, :)$ is known, then the AFS for $C(:, 2 : 4)$ is the intersection of the original AFS for C and a plane. The result is a bounded subset of an affine-linear space with two degrees of freedom. With a pre-given second spectrum the AFS for the concentration factor will be reduced by an additional degree of freedom.

6. Conclusion

The AFS appears to be a helpful tool for getting access to the range of all nonnegative factorizations of a given spectral data matrix. Inspecting the AFS of a system can support the user to get an idea on the possible solutions from which an MCR method can select the "one" final solution.

In this paper two objective functions are discussed which allow to decide whether a certain point belongs to the AFS or not. Further, the inverse polygon inflation algorithm has been introduced for the successful treatment of an AFS which consists of only one segment with a hole.

Mathematical proofs have been given on the boundedness of the AFS and on the fact that the origin is never contained in the AFS. The first property guarantees that numerical algorithms for the approximation of

the boundary of the AFS can successfully be used. The latter fact is the basis for the implementation of the inverse polygon inflation algorithm, which is to be used to compute an AFS with a hole. We observed such an AFS often for UV/Vis data. For FT-IR spectra an AFS with three separated segments seems to be typical.

There are still open questions on the AFS. The number of separated segments of an AFS appears to be unclear. Further fast and stable algorithms for the construction of the AFS for higher-dimensional systems are still to be developed.

Acknowledgement

The authors are grateful to the referees for their helpful comments. Section 4.6 on line-shaped AFS segments has been added after one referee has pointed out a problem with the original implementation of *FACPACK*.

References

- [1] W.H. Lawton and E.A. Sylvestre. Self modelling curve resolution. *Technometrics*, 13:617–633, 1971.
- [2] O.S. Borgen and B.R. Kowalski. An extension of the multivariate component-resolution method to three components. *Analytica Chimica Acta*, 174:1–26, 1985.

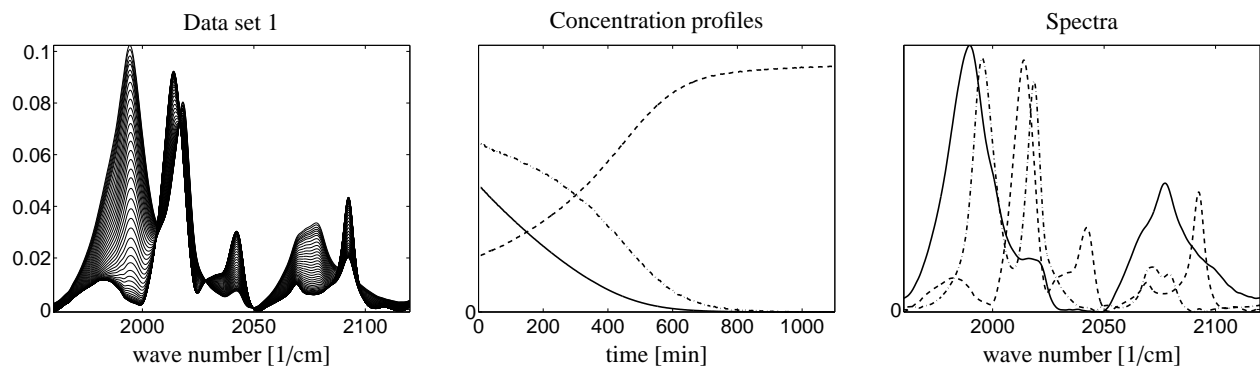


Figure 8: Data set 1 as described in Section 1.1. The solid line represents the olefin, the dash-dotted line stands for the acyl complex and the dashed line is for the hydrido complex.

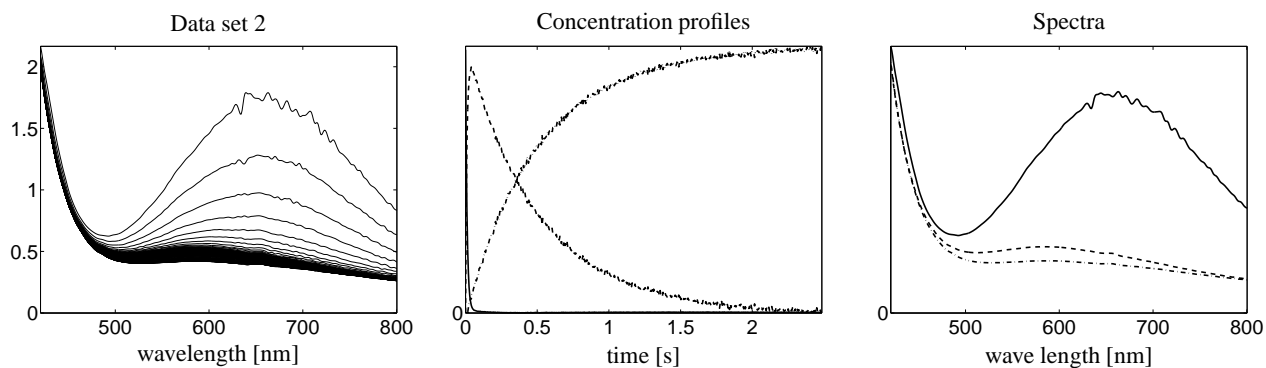


Figure 9: Data set 2 as described in Section 1.1 as well as the concentration profiles and spectra taken from [12].

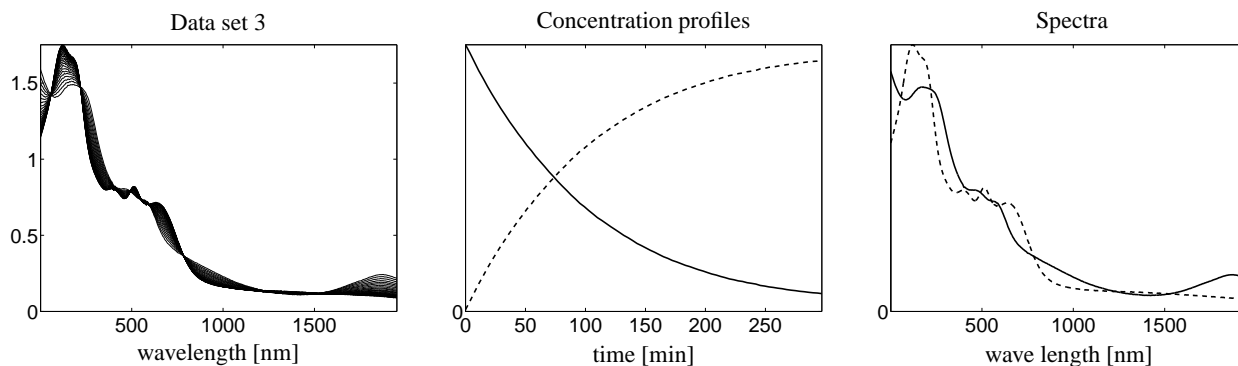


Figure 10: Data set 3 as described in Section 1.1 and the concentration profiles plus spectra taken from [13].

- [3] R. Rajkó and K. István. Analytical solution for determining feasible regions of self-modeling curve resolution (SMCR) method based on computational geometry. *J. Chemometrics*, 19(8):448–463, 2005.
- [4] M. Vosough, C. Mason, R. Tauler, M. Jalali-Heravi, and M. Maeder. On rotational ambiguity in model-free analyses of multivariate data. *J. Chemometrics*, 20(6-7):302–310, 2006.
- [5] H. Abdollahi, M. Maeder, and R. Tauler. Calculation and Meaning of Feasible Band Boundaries in Multivariate Curve Resolution of a Two-Component System. *Analytical Chemistry*, 81(6):2115–2122, 2009.
- [6] H. Abdollahi and R. Tauler. Uniqueness and rotation ambiguities in Multivariate Curve Resolution methods. *Chemometrics and Intelligent Laboratory Systems*, 108(2):100–111, 2011.
- [7] A. Golshan, H. Abdollahi, and M. Maeder. Resolution of Rotational Ambiguity for Three-Component Systems. *Analytical Chemistry*, 83(3):836–841, 2011.
- [8] 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. Chemometrics*, 27:106–116, 2013.
- [9] C. Kubis, D. Selent, M. Sawall, R. Ludwig, K. Neymeyr, W. Baumann, R. Franke, and A. Börner. Exploring between the extremes: Conversion dependent kinetics of phosphite-modified hydroformylation catalysis. *Chemistry - A European Journal*, 18(28):8780–8794, 2012.
- [10] T. Beweries, C. Fischer, S. Peitz, V. V. Burlakov, P. Arndt, W. Baumann, A. Spannenberg, D. Heller, and U. Rosenthal. Combination of spectroscopic methods: In situ NMR and UV/Vis measurements to understand the formation of group 4 metallocyclopentanes from the corresponding metallocyclopropenes. *Journal of the American Chemical Society*, 131(12):4463–4469, 2009.
- [11] C. Fischer, T. Beweries, A. Preetz, H.-J. Drexler, W. Baumann, S. Peitz, U. Rosenthal, and Heller D. Kinetic and mechanistic investigations in homogeneous catalysis using operando UV/Vis spectroscopy. *Catalysis Today*, 155(3–4):282–288, 2010.
- [12] 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. Chemometrics*, 26:526–537, 2012.
- [13] 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.
- [14] H. Minc. *Nonnegative matrices*. John Wiley & Sons, New York, 1988.
- [15] G.H. Golub and C.F. Van Loan. *Matrix computations*. Johns Hopkins University Press, Baltimore, MD, third edition, 1996.
- [16] J. Dennis, D. Gay, and R. Welsch. Algorithm 573: An adaptive nonlinear least-squares algorithm. *ACM Transactions on Mathematical Software*, 7:369–383, 1981.
- [17] M. Sawall and K. Neymeyr. Users’ guide to *FAC-PACK*. A software for the computation of multi-component factorizations and the area of feasible solutions. Software version 1.0. <http://www.math.uni-rostock.de/facpack/>, 2013.
- [18] 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. *Analytica Chimica Acta*, 791(0):25–35, 2013.
- [19] R. Rajkó. Additional knowledge for determining and interpreting feasible band boundaries in self-modeling/multivariate curve resolution of two-component systems. *Analytica Chimica Acta*, 661(2):129–132, 2010.
- [20] E.R. Malinowski. Window factor analysis: Theoretical derivation and application to flow injection analysis data. *J. Chemometrics*, 6(1):29–40, 1992.
- [21] J.C. Hamilton and P.J. Gemperline. Mixture analysis using factor analysis. II: Self-modeling curve resolution. *J. Chemometrics*, 4(1):1–13, 1990.
- [22] R. Manne. On the resolution problem in hyphenated chromatography. *Chemometrics and Intelligent Laboratory Systems*, 27(1):89–94, 1995.