

Simultaneous construction of dual Borgen plots.

II: Algorithmic enhancement for applications to noisy spectral data.

Mathias Sawall^a, Annekathrin Moog^a, Christoph Kubis^b, Henning Schröder^a, Detlef Selent^b, Robert Franke^{c,d},
Alexander Brächer^c, Armin Börner^b, Klaus Neymeyr^{a,b}

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

^bLeibniz-Institut für Katalyse e.V. an der Universität Rostock, Albert-Einstein-Straße 29a, 18059 Rostock

^cEvonik Performance Materials GmbH, Paul-Baumann Straße 1, 45772 Marl, Germany

^dLehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, 44780 Bochum, Germany

Abstract

Borgen plots are low-dimensional representations of the set of all nonnegative factorizations of spectral data matrices. Classical Borgen plots are limited to nonnegative data and can be constructed for the spectral factor or for the concentration profiles.

In the first part of this paper a simultaneous construction of the two dual Borgen plots is presented, which intensively exploits the underlying duality principles. The second part introduces algorithmic enhancements which make the simultaneous Borgen plot construction possible for noisy experimental data matrices which can contain small negative matrix entries. The new method is tested for FT-IR spectral data from the Rhodium catalyzed hydroformylation process. The results are compared to those by the *FACPACK*-implementation of the polygon inflation method.

Key words: multivariate curve resolution, nonnegative matrix factorization, Borgen plots, band boundaries of feasible solutions, polygon inflation, *FACPACK*.

1. Introduction

The topic of this paper is the simultaneous construction of pairs of dual Borgen plots. Borgen plots [1, 2, 3, 4] represent in a low-dimensional form the possible nonnegative factors of a spectral data matrix whose rows contain a series of spectra recorded from a chemical reaction system. For a detailed introduction to the underlying multivariate curve resolution (MCR) problem we refer to the first part of this paper [5] and the literature cited therein. See also [1, 2, 3, 6, 7, 8, 9, 10, 11] for an introduction and definition of the so-called Area of Feasible Solutions (AFS). Its geometric construction is developed by Borgen and Kowalski in their landmark paper from 1985 [1].

The construction of Borgen plots is inseparably linked to the existence of a nonnegative matrix factorization according to the underlying bilinear model. The first factor is the matrix of concentration profiles of the pure components and the second factor contains the pure component spectra. The nonnegativity cannot be guaranteed for experimental spectral data due to data preprocessing steps as background subtractions or baseline corrections. Another possible source of (small) negative matrix entries is a low-rank approximation of the spectral data matrix. Such a low rank approximation is often used by MCR methods, which employ the singular value decomposition (SVD) in order to reduce noise and the influence of nonlinearities or other perturbations. For such cases with relatively small negative matrix entries the concept of generalized Borgen plots (GBP) has been suggested in [3, 8]. In the GBP construction nonnegative convex combinations are substituted by affine combinations together with lower bounds (negative and close to zero) on the matrix entries of the pure component factors. Such affine combinations provide a sound mathematical basis for a generalization of Borgen plots. However, their robustness is smaller than that of the purely numerical approximation methods as the triangle enclosure algorithm [7] or the polygon inflation method [10]. Another difference is that the generalized Borgen plots work with absolute lower bounds for the negative entries whereas the numerical approaches use bounds on the relative size of negative entries. Here we suggest a new hybrid method for the simultaneous geometric construction of the dual Borgen plots which combines the speed and elegance of the geometric construction with the robustness and flexibility of the numerical approaches.

1.1. Idea and benefit of the hybrid method

The pairs of polygons FIRPOL and INNPOL each for the concentration factor and for the spectral factor are of key importance for the Borgen plot construction, see [1]. In [5] the simultaneous construction of dual Borgen plots first forms the two inner polygons INNPOL. Then the corresponding two FIRPOL polygons can easily be formed by duality relations, see also [12]. After these construction steps the triangle rotation algorithm is applied to the FIRPOL-INNPOL pair of polygons for the spectral factor and, simultaneously, the Borgen plot (or AFS) for the concentration factor is constructed by exploiting the duality relations.

The new hybrid algorithm is based on a similar approach. The idea is as follows:

- First compute certain enlarged supersets of the two polygons FIRPOL. These supersets are polygons which enclose the final AFS and also represent solutions with small negative entries. These polygons are determined by a variant of the polygon inflation algorithm [11].
- In a second step, the associated dual INNPOL polygons are constructed. The duality relates vertices of the superset of FIRPOL to bounding straight lines of the modified polygon INNPOL.
- In a third and final step the triangle rotation in combination with the duality principles is applied to the modified polygons FIRPOL and INNPOL in the way as described in the first part of this paper [5]. This results in the two desired approximate AFS-sets for the concentration profile factor and for the spectra factor.

The numerical results of the new hybrid method show a very good agreement with independently computed AFS approximations by means of the polygon inflation algorithm. The hybrid method is as fast as the method which is introduced in the first part of this paper. The method uses the duality principles not only to form INNPOL for given FIRPOL (in [5] the reverse construction is used from INNPOL to FIRPOL) but also for the simultaneous AFS construction. The second AFS-set is a “by-product” of the construction of the first AFS.

The benefit of the hybrid algorithm is as follows: The method improves the speed of the geometric AFS construction [1, 2, 3] and combines it with the robustness and flexibility for noisy/perturbed data which is typical for the numerical optimization-based methods as polygon inflation [10, 11], grid search [13, 14], triangle-chain boundary enclosure [7] or ray-casting [15]. The new algorithm computes the two dual AFS-sets with nearly the same speed as the classical Borgen plot algorithm computes a single AFS-set. The hybrid approach makes it possible to deal with slightly negative matrix entries.

1.2. Organization of the paper

Section 2 contains a brief overview on the relevant geometric objects as well as matrices and recapitulates the key results of the first part of this paper. The new hybrid construction method for noisy and perturbed data is introduced and explained in Section 3. Numerical results are presented in Section 4 for FT-IR spectra from the Rhodium catalyzed hydroformylation process. The results of the hybrid method are compared to those by the *FACPACK*-implementations of the polygon inflation method and the generalized Borgen plot algorithm [3].

1.3. Notation

The following variables are used in this paper. Variables with a tilde superscript refer to noisy data approximation.

D	$k \times n$ spectral data matrix by (1), (2).
C	$k \times s$ concentration matrix by (1), (2).
S	$n \times s$ spectral matrix by (1), (2).
$U\Sigma V^T$	singular value decomp. of D by (2).
T	$s \times s$ transformation by (2), (3).
$\mathcal{M}_C, \mathcal{M}_S$	AFS for concentration factor C by (5) or spectral factor by (4).
$\widetilde{\mathcal{M}}_C, \widetilde{\mathcal{M}}_S$	noisy data approximations of $\mathcal{M}_C, \mathcal{M}_S$ see Eqns. (12) and (11).
$\mathcal{I}_C, \mathcal{I}_S$	INNPOL for C, S by (9),(7).
$\widetilde{\mathcal{I}}_C, \widetilde{\mathcal{I}}_S$	noisy data approximations of $\mathcal{I}_C, \mathcal{I}_S$, see Eqns. (16) and (15).
$\mathcal{F}_C, \mathcal{F}_S$	FIRPOL for C, S by (9), (6).
$\widetilde{\mathcal{F}}_C, \widetilde{\mathcal{F}}_S$	noisy data approximations of $\mathcal{F}_C, \mathcal{F}_S$, see Eqns. (14) and (13).

2. A summary of dual Borgen plots

We give a brief summary on the dual Borgen plot construction from [5], see also [1, 2, 3, 7, 9, 10, 15, 16, 17]. The starting point is the Lambert-Beer model

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

with the $k \times n$ data matrix D , its $k \times s$ nonnegative matrix factor of pure concentration profiles and the $n \times s$ nonnegative matrix of the pure component spectra. The number of the pure components is s .

For given D we are interested in C and S . By means of a singular value decomposition (SVD) $D = U\Sigma V^T$ any of the desired factorizations of D can be written in the form

$$C = U\Sigma T^{-1}, \quad S^T = TV^T. \quad (2)$$

The matrix elements of the regular matrix $T \in \mathbb{R}^{s \times s}$ are the degrees of freedom. Only those T are relevant which result in nonnegative C and S . There are many nonnegative (or feasible) factorizations. This fact is known as rotational ambiguity of the decomposition. The area of feasible solutions (AFS), see e.g. [1, 2, 7, 10, 16] is a set of expansion coefficients of the columns of either C or S with respect to the basis of either left or right singular vectors of D . The expansion coefficient of the first left/right singular vector can be set equal to 1 (see [18] for a justification by the Perron-Frobenius theory) so that T can be written as

$$T = \begin{pmatrix} 1 & x_1 & \cdots & x_{s-1} \\ 1 & & & \\ \vdots & & W & \\ 1 & & & \end{pmatrix} \quad (3)$$

with an $(s-1) \times (s-1)$ regular matrix W . For given D with the rank s (noise-free case) the spectral AFS reads

$$\mathcal{M}_S = \{x \in \mathbb{R}^{s-1} : \text{exists } W \in \mathbb{R}^{(s-1) \times (s-1)} \text{ so that } \text{rank}(T) = s \text{ and } C, S \geq 0\}. \quad (4)$$

Analogously the AFS for the concentration factor reads

$$\mathcal{M}_C = \{y \in \mathbb{R}^{s-1} : \text{exists regular } T \in \mathbb{R}^{s \times s} \text{ with } (T^{-1})(:, 1) = (1, y^T)^T, C, S \geq 0\}. \quad (5)$$

The classical geometric construction [1] of \mathcal{M}_S for $(s=3)$ -component systems is based on the polygons FIRPOL \mathcal{F}_S and INNPOL \mathcal{I}_S . FIRPOL is defined as

$$\mathcal{F}_S = \left\{ x \in \mathbb{R}^{s-1} : V \begin{pmatrix} 1 \\ x \end{pmatrix} \geq 0 \right\}. \quad (6)$$

INNPOL is the convex hull

$$\mathcal{I}_S = \text{convhull}(\{a_i, i = 1, \dots, k\}). \quad (7)$$

with the vectors

$$a_i = \frac{(U\Sigma)^T(2 : s, i)}{(U\Sigma)^T(1, i)}, \quad i = 1, \dots, k, \quad (8)$$

which due to $DV = U\Sigma$ are the right singular vector representatives of the rows of D . Analogously, the sets \mathcal{F}_C and \mathcal{I}_C for factor C are

$$\mathcal{F}_C = \left\{ y \in \mathbb{R}^{s-1} : U\Sigma \begin{pmatrix} 1 \\ y \end{pmatrix} \geq 0 \right\}, \quad \mathcal{I}_C = \text{convhull}(\{b_j, j = 1, \dots, n\}) \quad (9)$$

with $b_j = (V^T(2 : s, j))/(V^T(1, j))$; see (7) in [5].

In [5] the dual Borgen plot construction rotates a tangent around \mathcal{I}_S in a way that for each tangent a triangle tightly including \mathcal{I}_S is constructed with one of its edges being located on the tangent. The third vertex of the triangle (which is not on the tangent) is used to form the inner boundary curve of the AFS. The dual Borgen plot construction uses only one rotation process around \mathcal{I}_S and results in the two AFS-sets \mathcal{M}_S and \mathcal{M}_C by exploiting duality principles.

3. Dual Borgen plots for noisy data

The dual Borgen plot method [5] cannot directly be applied to noisy/perturbed data with small negative data entries. Sometimes for such data INNPOL intersects the boundary of FIRPOL. Such an intersection necessarily occurs if the rank- s approximation of D has (small) negative matrix entries. Then at least one vector a^* of the vectors a_i by (8) represents a vector with at least one negative component. Thus a^* cannot be an element of FIRPOL by (6), since FIRPOL contains by its definition only the representatives of nonnegative vectors. As further \mathcal{I}_S is the convex hull of all a_i , the polygon INNPOL intersects the boundary of FIRPOL. Then the geometric construction of \mathcal{M}_S and \mathcal{M}_C is impossible as no triangle (case $s = 3$) exists which includes INNPOL and which is included in FIRPOL. Anticipating the model problem and results which are gained later in this paper, we refer to the two upper subplots of Fig. 3 that clearly demonstrate the intersection of INNPOL and FIRPOL. A possible and to some extent obvious approach is to start with increasing the size of FIRPOL. Such an enlargement can be justified: It is a known fact, see [3], that a weakening of the nonnegativity constraint (in a sense that small negative data values in the solution factors are accepted) enlarges the size of FIRPOL. These enlarged polygons for noisy data D are denoted by $\tilde{\mathcal{F}}_C$ and $\tilde{\mathcal{F}}_S$. They can easily be computed by the polygon inflation algorithm [10, 11]. The computational costs are low. Then in a second step the duality theory (also called a duality/complementarity theory), see e.g. [19, 20, 21, 22, 23], can be applied in order to compute the dual inner polygons $\tilde{\mathcal{I}}_C$ and $\tilde{\mathcal{I}}_S$. These polygons are consistent to the weakened nonnegativity constraint. The final step of a triangle rotation in $\tilde{\mathcal{F}}_S$ and around $\tilde{\mathcal{I}}_S$ is the same as suggested in [5]. Next all these steps are explained in detail.

3.1. Acceptance of negative data entries

The rank of noisy spectral data matrices D is larger than the number s of chemical components in the reaction system. Typically, there are s singular values which are characteristically larger than zero whereas the remaining singular values are close to zero. A rank- s approximation of D which is the starting point of the AFS computation can have negative entries. Other sources of negative entries of D are background subtractions or baseline corrections.

If D includes small negative entries, then one has to accept small negative entries of C and S as otherwise no nonnegative factorization of D exists. In the polygon inflation method for AFS computations [10, 11] the lower bound on the relative size of negative matrix entries is set to

$$\frac{C_{ij}}{\max |C(:, j)|} \geq -\varepsilon_C, \quad i = 1, \dots, k, \quad \text{and} \quad \frac{S_{\ell j}}{\max |S(:, j)|} \geq -\varepsilon_S, \quad \ell = 1, \dots, n, \quad (10)$$

for all $j = 1, \dots, s$ and with control parameters $\varepsilon_C, \varepsilon_S \geq 0$.

Alternatively, the size of negative matrix entries can be controlled by *ssq*-bounds, see [7, 13, 14, 24, 25], or by using absolute lower limits which accept only solutions with

$$C_{ij} \geq -\tilde{\varepsilon}_C, \quad S_{\ell j} \geq -\tilde{\varepsilon}_S$$

for proper control parameters $\tilde{\varepsilon}_C, \tilde{\varepsilon}_S \geq 0$. Absolute bounds are used in the generalized Borgen plots construction [3, 8].

Here, we prefer to use the relative bounds (10). On this basis we use in the following generalizations of the AFS-sets

$$\mathcal{M}_S = \left\{ x \in \mathbb{R}^{s-1} : \text{exists } W \in \mathbb{R}^{(s-1) \times (s-1)} \text{ so that } \text{rank}(T) = s, C, S \text{ fulfill (10)} \right\} \quad (11)$$

with T as in (3) and C and S as in (2) and

$$\mathcal{M}_C = \left\{ y \in \mathbb{R}^{s-1} : \text{exists } T \in \mathbb{R}^{s \times s} \text{ with } \text{rank}(T) = s, (T^{-1})(:, 1) = (1, y^T)^T, C, S \text{ fulfill (10)} \right\}. \quad (12)$$

3.2. Generalization of \mathcal{F}_S and \mathcal{F}_C for noisy data

FIRPOL, according to (6), is the set of all x for which the linear combination $V(1, x^T)^T$ is componentwise non-negative. With respect to the relative lower bound (10) for $j = 1$ the generalized set FIRPOL is defined as

$$\tilde{\mathcal{F}}_S = \left\{ x \in \mathbb{R}^{s-1} : \min_{i=1, \dots, n} \frac{(1, x^T)V(i, :)^T}{\|(1, x^T)V^T\|_\infty} \geq -\varepsilon_S \right\}. \quad (13)$$

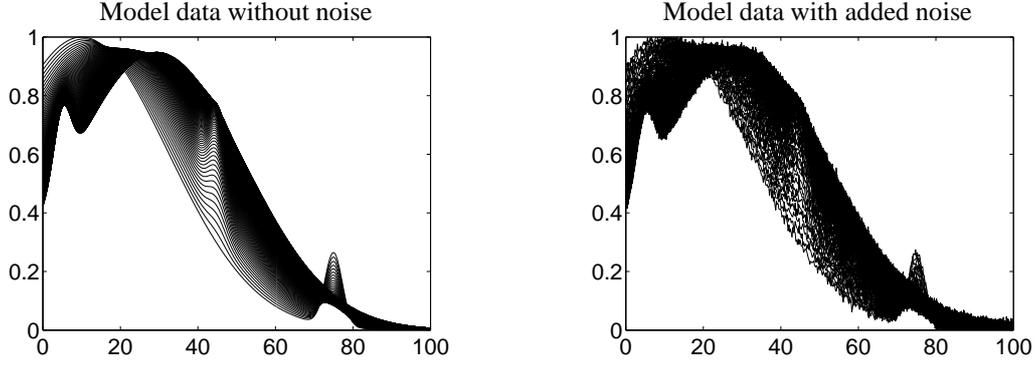


Figure 1: Model data with and without noise of a three-component system in order to illustrate the polygons $\mathcal{F}_C, \mathcal{F}_S, \mathcal{I}_C$ and \mathcal{I}_S and their pendants for noisy data. Left: The model data set from [5] for $k = 1000, n = 500$. Right: The data after application of normal distributed pseudo random noise with mean zero and a standard deviation of 0.01.

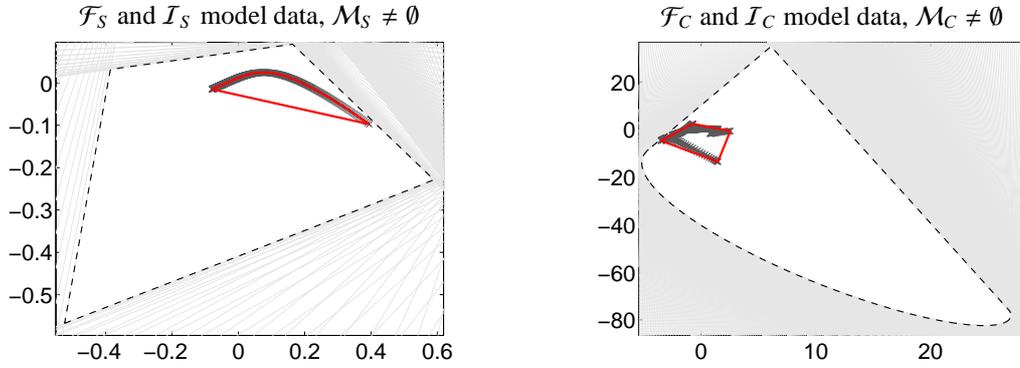


Figure 2: Left plot: Polygons \mathcal{F}_S (dashed) and \mathcal{I}_S (red). Right plot: \mathcal{F}_C (dashed) and \mathcal{I}_C (red). The polygons are plotted for the model data shown in the left subplot of Fig. 1. In the polygons INNPOL the data representatives a_i and b_i are plotted in gray color. The AFS-sets \mathcal{M}_S and \mathcal{M}_C are nonempty.

Analogously the generalization of \mathcal{F}_C with respect to (10) for $j = 1$ reads

$$\tilde{\mathcal{F}}_C = \left\{ y \in \mathbb{R}^{s-1} : \min_{i=1, \dots, k} \frac{U(i, \cdot) \Sigma(1, y^T)^T}{\|U \Sigma(1, y^T)^T\|_\infty} \geq -\varepsilon_C \right\}. \quad (14)$$

3.3. Computation of $\tilde{\mathcal{F}}_S$ and $\tilde{\mathcal{F}}_C$

The two polygons $\tilde{\mathcal{F}}_S$ and $\tilde{\mathcal{F}}_C$ can easily and precisely be computed by means of the polygon inflation algorithm (PIA). The idea of PIA is to approximate the boundary of a bounded set by a sequence of adaptively refined polygons. The vertices of the polygons are located on the surface of the set. The boundary of the set is assumed to be locally representable by continuous functions. In [11] the PIA algorithm has been designed for AFS computations. Here PIA is used for the approximation of the polygons $\tilde{\mathcal{F}}_S$ and $\tilde{\mathcal{F}}_C$. This makes the computation very easy. In contrast to [11] no complicated and computationally expensive target function must be evaluated. Instead, according to (6) and (9) a point x (resp. y) belongs to the surface of the FIRPOL polygons if $V(1, x)^T \geq 0$ (resp. $U \Sigma(1, y)^T \geq 0$) and $V(1, \eta x)^T$ (resp. $U \Sigma(1, \eta y)$) contains at least one negative component for $\eta > 1$. More generally, for the noisy-data case the zero transition for changing η turns into a $-\varepsilon_S$ or $-\varepsilon_C$ transition. Again, an adaptive polygon refinement strategy enables a highly accurate boundary approximation. All this makes the computational costs low.

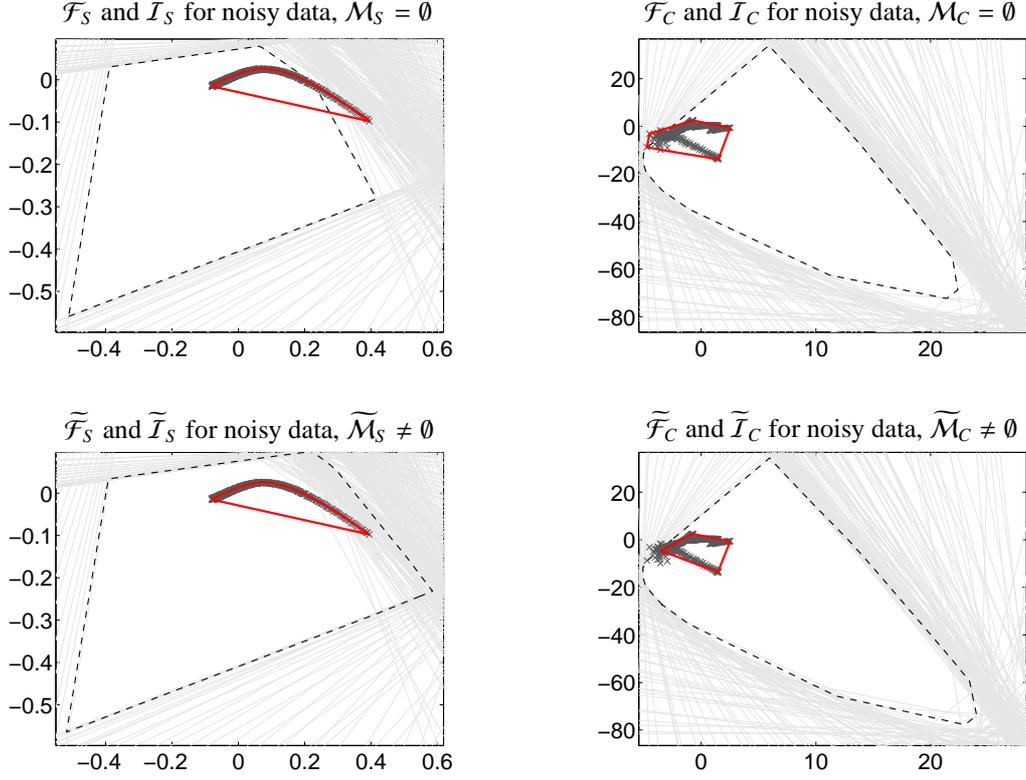


Figure 3: The polygons FIRPOL and INNPOL are drawn for the data set with noise as shown in the right subplot of Fig. 1. The two upper subplots demonstrate that INNPOL \mathcal{I} (red) is not a subset of FIRPOL \mathcal{F} (dashed) for these data. Hence the AFS is empty, i.e. no nonnegative factorization of D exists. The two lower plots show the generalized polygons $\tilde{\mathcal{F}}$ and $\tilde{\mathcal{I}}$ each for the factors C and S . The AFS-sets \mathcal{M}_S and \mathcal{M}_C are nonempty. The boundary lines to strict nonnegativity with $x_1 V_{j2} + x_2 V_{j3} = -V_{j1}$, $j = 1, \dots, n$, resp. $y_1 \sigma_2 U_{i2} + y_2 \sigma_2 U_{i3} = -\sigma_1 U_{i1}$, $i = 1, \dots, k$, are plotted in faint gray color. The data representing points a_i , $i = 1, \dots, k$ resp. b_j , $j = 1, \dots, n$ are plotted in the INNPOL polygons as by gray crosses.

3.4. Generalization of INNPOL for noisy data

The duality theory in [5, 12, 19, 20] describes how a point on the boundary of the polygon \mathcal{F}_C relates to an affine hyperplane which touches the dual inner polygon \mathcal{I}_S . Further, Theorem 3.6 of [5] proves that a vertex of \mathcal{F}_C relates to a (true) facet of \mathcal{I}_S . Our aim is to use these duality relations in order to construct from the boundary of the generalized set FIRPOL $\tilde{\mathcal{F}}_C$ the dual set $\tilde{\mathcal{I}}_S$. Hence $\tilde{\mathcal{I}}_S$ is constructed by exploiting the duality relations. A direct construction by taking the convex hull of the vectors a_i , $i = 1, \dots, k$, only results in \mathcal{I}_S and is not consistent with the weakened nonnegativity constraint.

The construction of $\tilde{\mathcal{I}}_S$ is based on its duality to $\tilde{\mathcal{F}}_C$. The duality relation of a point z to its dual hyperplane $E = \{y \in \mathbb{R}^{s-1} : y^T z = -1\}$ is introduced in Def. 3.1 in [5]. In Lemma 3.3 and Theorem 3.5 of [5] the focus is on duality of boundary points of $\tilde{\mathcal{F}}_C$ to tangential hyperplanes of $\tilde{\mathcal{I}}_S$ in order to construct $\tilde{\mathcal{F}}_C$ from $\tilde{\mathcal{I}}_S$. However, the duality of polygons is an involutory mapping which relates in the same way the boundary of $\tilde{\mathcal{I}}_S$ to tangential hyperplanes of $\tilde{\mathcal{F}}_C$. Thus the polygon $\tilde{\mathcal{I}}_S$ can be defined as

$$\tilde{\mathcal{I}}_S = \{x \in \mathbb{R}^{s-1} : \text{the dual hyperplane } E \text{ of } x \text{ has an empty intersection with the interior of } \tilde{\mathcal{F}}_C\}. \quad (15)$$

Equivalently, the empty intersection of E with the interior of $\tilde{\mathcal{F}}_C$ can be expressed in a way that E has either an empty intersection with $\tilde{\mathcal{F}}_C$ or at most touches its boundary. An alternative way to write this polygon is

$$\tilde{\mathcal{I}}_S = \text{closure}(\{x \in \mathbb{R}^{s-1} : E \cap \tilde{\mathcal{F}}_C = \emptyset \text{ for } E \text{ being the dual affine hyperplane of } x\}),$$

where closure denotes the topological closure of a set, i.e. the set augmented by all its limit points. Analogously, the set

$$\widetilde{\mathcal{I}}_C = \left\{ y \in \mathbb{R}^{s-1} : \text{the dual hyperplane } E \text{ of } y \text{ has an empty intersection with the interior of } \widetilde{\mathcal{F}}_S \right\} \quad (16)$$

is the pendant of the polygon \mathcal{I}_C for the case of noisy data.

3.5. Computation of $\widetilde{\mathcal{I}}_S$ from $\widetilde{\mathcal{F}}_C$

Numerically, we compute $\widetilde{\mathcal{I}}_S$ by constructing the dual points of the edges of the convex hull of $\widetilde{\mathcal{F}}_C$. These relations are illustrated in Figs. 1 and 2 in [5]. Alternatively, one can compute for the vertices of the convex hull of $\widetilde{\mathcal{F}}_C$ the associated dual half-spaces which are oriented in a way that each includes the origin; see Lemma 3.5 in [5]. The intersection of all these half-spaces yields $\widetilde{\mathcal{I}}_S$ (and is close to the initial polygon \mathcal{I}_S).

The figures 1, 2 and 3 illustrate the polygons \mathcal{I}_S and \mathcal{I}_C and their noisy data pendants for the three-component model problem described in [5]. The dimensions are $k = 1000$ and $n = 500$. First Fig. 1 shows (a subset of) the series of spectra without and with noise. We have added normal distributed pseudo random noise with mean zero and a standard deviation of 0.01. It holds for the relative componentwise minimum/maximum that $\min(D)/\max(D) = -3.3 \cdot 10^{-2}$ and for the rank-3 approximation that

$$\frac{\|D - U\Sigma(:, 1:3)V(:, 1:3)^T\|_F}{\|D\|_F} = 1.6644 \cdot 10^{-2}. \quad (17)$$

For the noise-free case Figure 2 illustrates the polygons \mathcal{F}_S and \mathcal{I}_S as well as \mathcal{F}_C and \mathcal{I}_C together with their generating boundary lines (around \mathcal{F}_C) and generating points (the gray stars in \mathcal{I}_S). Finally, Fig. 3 shows the situation for noisy data. First the polygons are drawn without any consideration for the perturbed data. Then no triangle exists in FIRPOL which includes INNPOL. Consequently the AFS-sets \mathcal{M}_S and \mathcal{M}_C are empty. Further the generalized noise-adapted polygons $\widetilde{\mathcal{F}}_S$ and $\widetilde{\mathcal{I}}_S$ resp. $\widetilde{\mathcal{F}}_C$ and $\widetilde{\mathcal{I}}_C$ are plotted. For these polygons the AFS-sets $\widetilde{\mathcal{M}}_S$ and $\widetilde{\mathcal{M}}_C$ are nonempty. We have to show that these sets are good approximations of \mathcal{M}_S and \mathcal{M}_C for the case of noisy/perturbed data, see Sec. 4.5.

Remark 3.1. 1. *The computational costs for a numerical approximation of $\widetilde{\mathcal{F}}_S$ and $\widetilde{\mathcal{F}}_C$ by the polygon inflation algorithm are small, see Sec. 3.3.*

2. *For the computation of $\widetilde{\mathcal{F}}_C$ a precise boundary approximation is recommended in order to guarantee a stable subsequent application of the duality principles for the computation of $\widetilde{\mathcal{I}}_S$. The control parameter ε_b on the boundary precision is explained in Sec. 3.9.1 of [10]. A reasonable parameter choice is $\varepsilon_b \leq 10^{-6}$.*
3. *The control parameter ε_C limits the relative magnitude of negative entries of C according to (10); the parameter ε_S works similarly for the factor S . Increasing the value of ε_S results in an expansion of $\widetilde{\mathcal{F}}_S$ and thus in a shrinkage of the dual polygon $\widetilde{\mathcal{I}}_C$. (The shrinkage of $\widetilde{\mathcal{I}}_C$ is a consequence of Lemma 3.3 in [5] as the inner points of $\widetilde{\mathcal{I}}_C$ are dual to hyperplanes which do not contribute to the boundary of $\widetilde{\mathcal{F}}_S$ and vice versa.) The geometric construction process shows that $\widetilde{\mathcal{M}}_S$ is expanded at the outer boundary and that $\widetilde{\mathcal{M}}_C$ is expanded at its inner boundary. Similarly, increasing the value of ε_C results in an expansion of $\widetilde{\mathcal{F}}_C$ and in a shrinkage of the dual polygon $\widetilde{\mathcal{I}}_S$. Thus the set $\widetilde{\mathcal{M}}_S$ grows at its inner boundary and $\widetilde{\mathcal{M}}_C$ grows at the outer boundary. These relations are illustrated for example by Figure 6 in [3], Figure 17 in [9] and Figure 15 in [8].*

For noisy data ε_C and ε_S should not be chosen too small as otherwise no feasible factorizations exist that meet the conditions. Then the AFS is empty. It is an interesting question which parameter selections can yield isolated feasible solutions; see [26] for the so-called data-based uniqueness in the noisy data case.

We cannot recommend a fixed rule how to determine the values of ε_S and ε_C . The FACKPACK software always tries to use small parameter values. The polygon inflation algorithm first tries to compute a strict-nonnegative matrix factorization (NMF) of D in a way that the smallest entries of its factors are equal to or greater than 0. However, if no initial strict NMF can be found in this optimization process, then small positive values of ε_S and ε_C must be accepted. In other words, small matrix elements close to zero are accepted. If in such a case these parameters are equal to the smallest possible values so that an NMF exists, then at least one subset of the AFS is often very small. Then it cannot be guaranteed that the AFS includes the chemically correct solution. In this work we use $\varepsilon_S = \varepsilon_C = 5 \cdot 10^{-2}$ for the model problem from [5] which corresponds to the value of the rank-3 approximation error (17).

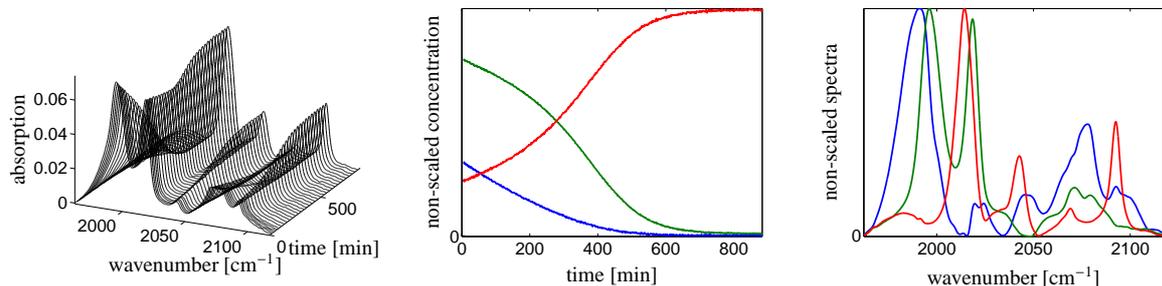


Figure 4: The experimental hydroformylation data set together with the pure component factors. Left: Every 25th spectrum of the 850 spectra is plotted. Center: The true concentration profiles computed by kinetic modeling. Right: The true pure component spectra. Color coding: Olefin (blue), acyl complex (green), hydrido complex (red).

3.6. Geometric AFS construction

Finally, the boundary curves of the AFS-sets \mathcal{M}_S and \mathcal{M}_C are determined. This is done by the simultaneous geometric construction algorithm, see the first part of this paper [5]. The only difference is that the algorithm is applied to $\tilde{\mathcal{F}}_S$, $\tilde{\mathcal{F}}_C$ and $\tilde{\mathcal{I}}_S$ instead of \mathcal{F}_S , \mathcal{F}_C and \mathcal{I}_S . The algorithm yields the inner parts of the boundary curves of the two AFS-sets. The outer parts of the boundary curves are given by parts of the boundaries of $\tilde{\mathcal{F}}_S$ and $\tilde{\mathcal{F}}_C$. The construction of the polygons $\tilde{\mathcal{M}}_C$ and $\tilde{\mathcal{M}}_S$ is based on the inequalities (10). The same constraints are used in order to compute the FIRPOL polygons by the polygon inflation algorithm. Thus the inner and outer parts of the boundary are fitting to one another for the case of noisy data.

3.7. Non-convexity of $\tilde{\mathcal{F}}_S$

The acceptance of small negative entries of C and S may have the consequence that the polygons $\tilde{\mathcal{F}}_S$ and $\tilde{\mathcal{F}}_C$ lose their convexity. Non-convexity is an unwanted side effect. On the one hand the non-convexity of $\tilde{\mathcal{F}}_C$ results in problems in the duality-based construction of $\tilde{\mathcal{I}}_S$. On the other hand the triangle rotation process for the simultaneous AFS construction, which determines the points of intersection of tangents of $\tilde{\mathcal{I}}_S$ with boundary of the surrounding polygon $\tilde{\mathcal{F}}_S$, can find more than two points of intersection in the case of a non-convex set $\tilde{\mathcal{F}}_S$. Then additional steps are required in the program implementation in order to select the correct points of intersection. In our program implementation we avoid such problems with “slightly” non-convex sets in a way that we always work with the convex hulls of $\tilde{\mathcal{F}}$ resp. $\tilde{\mathcal{F}}_C$.

4. Numerical results

The new algorithm is tested for an FT-IR experimental data set from the hydroformylation process. For this noise-including data set we first draw the polygons $\tilde{\mathcal{F}}$, $\tilde{\mathcal{F}}_C$ and $\tilde{\mathcal{I}}_S$. The series of spectra has undergone a background subtraction and a baseline correction. After this, the simultaneous geometric construction yields the AFS-sets \mathcal{M}_S and \mathcal{M}_C . These AFS-sets are compared to the corresponding approximations by the *FACPACK* software. Further the results are compared to the generalized Borgen plots. The results are analyzed and the computation times are discussed. All computations were run on a PC with an Intel 3.40GHz quad-core processor and 16GB RAM. The main algorithm is implemented in fast C-routines. The program is executed on a single core of the processor.

4.1. Rhodium catalyzed hydroformylation

From the series of 850 FT-IR spectra only a spectral window with 645 spectral channels is considered. Thus $D \in \mathbb{R}^{850 \times 645}$. The associated time and wavenumber intervals are $t \in [4.7, 883.6]\text{min}$ and $\nu \in [1962.3, 2117.6]\text{cm}^{-1}$. Three major absorbing components contribute to the spectra, namely an olefin component, an acyl complex and a hydrido complex. The reaction product, an aldehyde, shows no absorption in the spectral window. For the experimental details see [27]. The series of spectra and the true pure component spectra and concentration profiles, which are computed by using additionally a kinetic model, are presented in Fig. 4.

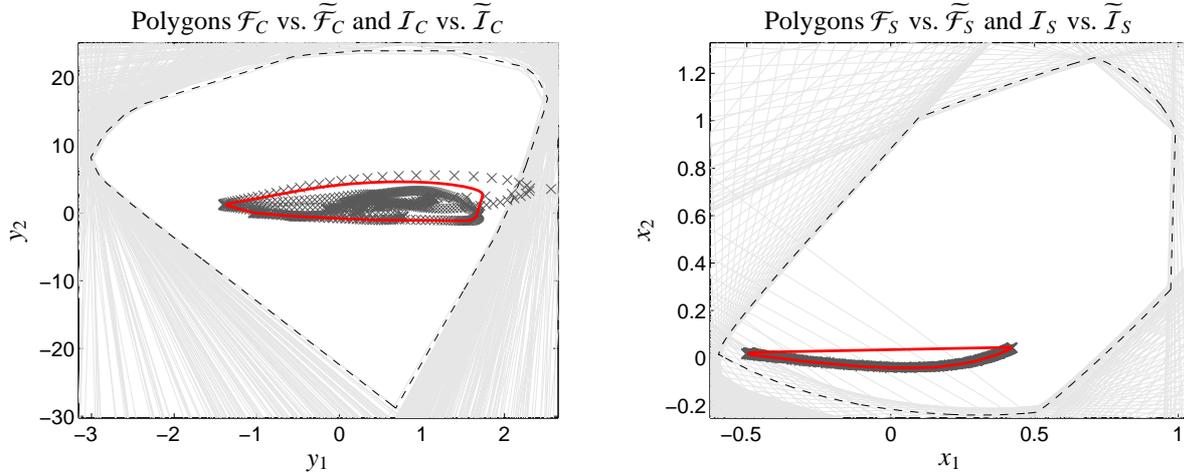


Figure 5: Comparison of the polygons \mathcal{F} and \mathcal{I} and their pendants for partially negative data $\tilde{\mathcal{F}}$ and $\tilde{\mathcal{I}}$ each for C and S . Left: The polygons $\tilde{\mathcal{F}}_C$ (black broken line) and $\tilde{\mathcal{I}}_C$ (red line). The gray lines define the boundary of \mathcal{F}_C and the dark gray crosses are the data representing vectors b_j , $j = 1, \dots, n$. Right: The polygons $\tilde{\mathcal{F}}_S$ (black broken line) and $\tilde{\mathcal{I}}_S$ (red line) are drawn. The gray lines define the boundary of \mathcal{F}_S and intersect in the left lower region the polygon $\tilde{\mathcal{F}}_S$. The dark gray crosses represent the data vectors a_i , $i = 1, \dots, k$. The left plot illustrates major differences between $\tilde{\mathcal{I}}_C$ and \mathcal{I}_C . The triangle rotation algorithm applied to \mathcal{F}_C and \mathcal{I}_C would result in an empty AFS \mathcal{M}_C . A similar situation is illustrated in the right plot, where \mathcal{I}_S intersects \mathcal{F}_S . In particular two half-spaces (by the gray lines) in the right plot do not include the origin due to the fact that $\min(D) = -4.8 \cdot 10^{-4}$.

Due to a background subtraction and also due to a subsequent SVD-based low-rank approximation the matrix D includes small negative entries. The smallest matrix entry of D is $\min(D) = -4.8 \cdot 10^{-4}$. The relative size of the smallest matrix entry equals $\min(D)/\max(D) = -6.5 \cdot 10^{-3}$; the minimum and the maximum are taken componentwise. Further, the relative error of the rank-3 approximation of D reads

$$\frac{\|D - CS^T\|_F}{\|D\|_F} = \sqrt{\frac{\sum_{i=4}^n \sigma_i^2}{\sum_{i=1}^n \sigma_i^2}} = 5.995 \cdot 10^{-3} \quad (18)$$

with C and S containing columnwise expansions with respect to the first $s = 3$ left- and right singular vectors. The “true” pure components are computed by adding a kinetic Michaelis-Menten hard model, cf. [28, 29, 30]. The ssq-deviations of the Michaelis-Menten solutions $C^{(\text{ode})}$ from the pure component factor are computed as

$$d_i = \frac{\|C^{(\text{ode})}(:, i) - \alpha(U\Sigma(1, y_1(i), y_2(i))^T)\|_2^2}{\|C^{(\text{ode})}(:, i)\|_2^2}$$

with norm-minimizing scaling constants α . The numerical values are $(2.2 \cdot 10^{-4}, 5.2 \cdot 10^{-5}, 1.5 \cdot 10^{-4})$ for the three components.

4.2. FIRPOL, INNPOL and parameter setting

First the polygons $\tilde{\mathcal{F}}_S$ and $\tilde{\mathcal{F}}_C$ are computed by polygon inflation with the control parameters $\varepsilon = \varepsilon_S = \varepsilon_C = 6.0 \cdot 10^{-3}$ as lower bounds for the relative sizes of negative entries in C and S , see Eqns. (13) and (14). See [10, 11] for the details on these control parameters. Then the duality theory in [5] is the basis to compute from $\tilde{\mathcal{F}}_C$ the inner polygon $\tilde{\mathcal{I}}_S$ for the case of noisy data, see Sec. 3.5.

Figure 5 shows the FIRPOL and INNPOL polygons for the hydroformylation data set together with their pendants for slightly negative data. Additionally, the spectral data representing vectors a_i , $i = 1, \dots, k$, and the b_j , $j = 1, \dots, n$ are plotted. The gray straight lines represent the affine half-spaces whose intersections define the FIRPOL polygons \mathcal{F}_S and \mathcal{F}_C . The left subplot in Figure 5 highlights that the data representatives and thus their convex hull INNPOL intersect FIRPOL. The consequence is an empty AFS. However, the smaller and modified set $\tilde{\mathcal{I}}_C$ (in red) is completely

control par.		number of vertices			computation times [s]	
ε_b	δ	$\tilde{\mathcal{F}}_S$	$\tilde{\mathcal{F}}_C$	$\tilde{\mathcal{I}}_S$	$\tilde{\mathcal{F}}_S$ & $\tilde{\mathcal{F}}_C$	$\tilde{\mathcal{I}}_S$
10^{-5}	10^{-4}	326	512	119	0.192	0.164
10^{-6}	10^{-5}	814	786	90	0.381	0.654
10^{-7}	10^{-6}	1 456	1 118	104	0.796	1.704
10^{-8}	10^{-7}	2 232	1 458	105	1.112	3.691
10^{-9}	10^{-8}	3 094	1 848	281	1.837	7.514

Table 1: Computation times for $\tilde{\mathcal{F}}_S$ and $\tilde{\mathcal{F}}_C$ and for the duality based construction of $\tilde{\mathcal{I}}_S$ from $\tilde{\mathcal{F}}_C$ for different precisions ε_b and different stopping criteria δ of polygon inflation. The parameter ε_b controls the accuracy of the boundary approximation. The stopping criteria δ controls the number of edges, see [10, 11, 31]. We use $\delta = 10\varepsilon_b$.

number of tangents		contr. param.		number of vertices		computation times [s] for $\tilde{\mathcal{M}}_S$ and $\tilde{\mathcal{M}}_C$		
m_{init}	m_{all}	ε_b	δ	$\tilde{\mathcal{M}}_S$	$\tilde{\mathcal{M}}_C$	$\tilde{\mathcal{F}}_S, \tilde{\mathcal{F}}_C, \tilde{\mathcal{I}}_S$	geom. constr.	Total comp. time
1 500	3 206	10^{-6}	10^{-5}	3 445	6 091	1.035	0.143	1.172
2 500	5 495	10^{-7}	10^{-6}	5 982	10 267	2.500	0.378	2.878
5 000	7 995	10^{-7}	10^{-6}	7 964	14 229	2.500	0.570	3.070
10 000	14 530	10^{-8}	10^{-7}	14 007	25 184	4.803	1.348	6.151
15 000	19 530	10^{-8}	10^{-7}	17 968	33 105	4.803	2.003	6.806

Table 2: The computation times of the dual Borgen plot method are listed for different numbers of tangents and different precision parameters ε_b . The computation times for computing the polygons $\tilde{\mathcal{F}}_S$, $\tilde{\mathcal{F}}_C$ and $\tilde{\mathcal{I}}_S$ as well as the times for the construction of the two dual AFS-sets are tabulated separately. The parameter setting $m_{\text{init}} = 2500$ and $\varepsilon_b = 10^{-7}$ is sufficient for a high-resolution computation.

contained in $\tilde{\mathcal{F}}_C$. The right subplot of Figure 5 shows two gray straight lines which bound two half-spaces which do not include the origin. By the mathematical theory this is only possible if $D^T D$ is a reducible matrix or if D includes negative entries. Here it holds componentwise that $\min(D) = -4.8 \cdot 10^{-4}$. In particular the first right singular vector contains positive and also two small negative entries (whereas for $D \geq 0$ the Perron-Frobenius theory guarantees that the first singular vector can be selected nonnegative). By the modified definition of $\tilde{\mathcal{F}}_S$ this unwanted effect disappears.

With the parameters $\varepsilon_b = 10^{-7}$ and $\delta = 10^{-6}$, see the *FACPACK* manual, the polygon $\tilde{\mathcal{F}}_S$ has 1456 vertices and $\tilde{\mathcal{F}}_C$ has 1118 vertices. The duality theory allows us to construct $\tilde{\mathcal{I}}_S$ from given $\tilde{\mathcal{F}}_C$. Then $\tilde{\mathcal{I}}_S$ has only 104 vertices. The much smaller number of vertices of $\tilde{\mathcal{I}}_S$ is due to the fact that the convex hull of $\tilde{\mathcal{F}}_C$ is used for the dual construction, see Sec. 3.7. This has characteristically reduced the number of vertices. Finally, the inner polygon $\tilde{\mathcal{I}}_C$ has 236 vertices. Figure 5 shows the polygons for a fixed set of control parameters. Table 1 lists the computation times for other parameter selections. The computation is still very fast for ε_b and δ close to zero, which results in a larger number of vertices of the polygons.

4.3. Geometric construction of the dual AFS-sets

The final step is the simultaneous geometric construction of the inner boundaries of the AFS-sets \mathcal{M}_S and \mathcal{M}_C . The construction algorithm is explained in the first part of this paper [5]. The only difference is that \mathcal{F}_S , \mathcal{F}_C and \mathcal{I}_S are substituted by $\tilde{\mathcal{F}}_S$, $\tilde{\mathcal{F}}_C$ and $\tilde{\mathcal{I}}_S$. We use the control parameters $\varepsilon_b = 10^{-7}$ and $\delta = 10^{-6}$. The resulting AFS-sets are shown in Fig. 6. These AFS-sets are plotted together with the approximations by the polygon inflation based AFS computation by *FACPACK* and also with the results by the generalized Borgen plots algorithm. The first two approximations coincide, but there are small differences to the generalized Borgen plots, see Fig. 6. The “true solution” (by kinetic hard-modeling) are marked by bold colored crosses. The dual Borgen plot uses a rotation angle increment of $360^\circ/5000 = 0.072^\circ$ and some additional tangents, especially those which cover edges of $\tilde{\mathcal{I}}_S$. The total number of used tangents is 7995. The geometric construction by duality takes 0.570s and the overall dual Borgen plot construction takes 3.070s.

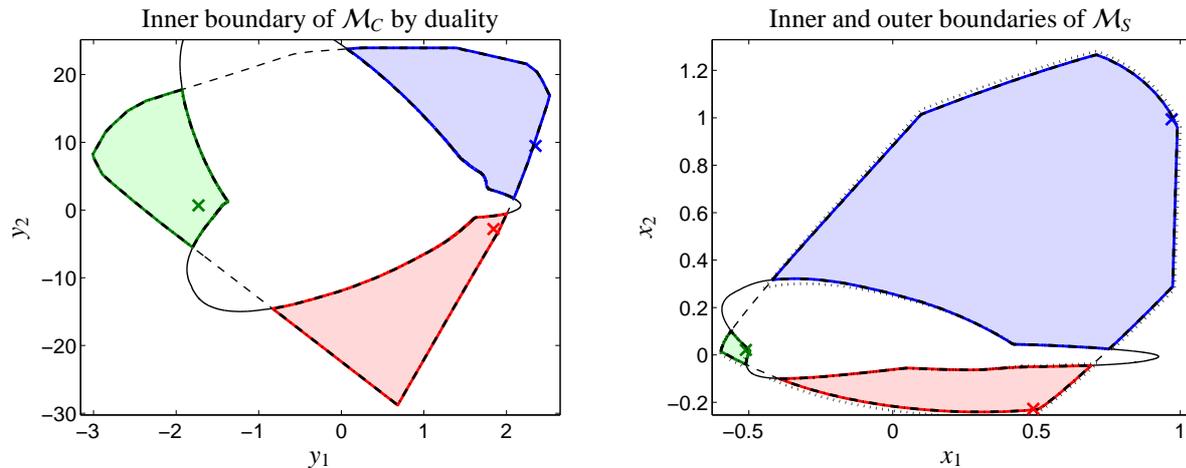


Figure 6: The AFS-sets \mathcal{M}_C and \mathcal{M}_S for the experimental spectral data set. The three isolated subsets for each of the two AFS-sets by the dual Borgen plot method are drawn in green, red and blue. The colored x-markers represent the Michaelis-Menten hard modeling results as plotted in Fig. 4. The residuals for all solutions in the AFS has the fixed value $5.995 \cdot 10^{-3}$ by Eq. (18). Their boundaries coincide with the dash-dotted lines which are the boundaries of the AFS approximations by the polygon inflation based AFS computation with the *FACPACK* software. Dotted black lines: The right figure also shows the AFS approximation by the generalized Borgen plot algorithm. Minor differences can be seen at the boundary of the blue set. Additionally the sets $\tilde{\mathcal{F}}$ and $\tilde{\mathcal{F}}_C$ are marked by broken black lines. The inner boundaries of the AFS-sets are plotted in black.

contr. param.		number of vertices		computation times [s] <i>FACPACK</i> -PIA		
ε_b	δ	$\tilde{\mathcal{M}}_S$	$\tilde{\mathcal{M}}_C$	$\tilde{\mathcal{M}}_S$	$\tilde{\mathcal{M}}_C$	$\tilde{\mathcal{M}}_S$ and $\tilde{\mathcal{M}}_C$
10^{-3}	10^{-3}	261	295	8.725	13.019	21.744
10^{-4}	10^{-3}	271	299	12.389	17.033	29.422
10^{-4}	10^{-4}	679	663	27.055	30.605	57.660
10^{-5}	10^{-4}	669	663	37.768	39.743	77.511

Table 3: The computation times are tabulated of the *FACPACK*-based AFS computation by the polygon inflation algorithm (*FACPACK*-PIA). The times are considerably larger than those of the dual Borgen plot construction, see Table 2. The boundaries of the AFS-sets by the two methods coincide apart from small deviations of the size of the control parameter ε_b .

4.4. Variation of the number of tangents

The dual Borgen plot construction is run with different numbers of tangents. For the computations we use initial numbers m_{init} of 1500, 2500, 5000, 10000 and 15000 tangents; consecutive tangents enclose a fixed rotation angle. These tangents are augmented to a total number m_{all} of tangents. These additional tangents primarily refer to edges of $\tilde{\mathcal{I}}_S$ by the duality relations. The computation times for the polygon constructions and for the following duality-based simultaneous AFS construction are listed in Table 2. The computation times remain to be in the range of few seconds even for the highest resolution. These results are to be compared with alternative methods for AFS computations.

4.5. Comparison of the results

The AFS-sets $\tilde{\mathcal{M}}_S$ and $\tilde{\mathcal{M}}_C$ for perturbed/noisy data are compared to *FACPACK* based AFS computation by the polygon inflation algorithm (*FACPACK*-PIA) and also to the generalized Borgen plots (GBP). For these comparisons we consider different choices of the control parameters.

4.5.1. Comparison to polygon inflation method

Table 3 contains the computing times for the AFS computations by *FACPACK*-PIA for different settings of the boundary precision ε_b and the stopping criteria δ . Even for the default control parameter settings $\varepsilon_b = \delta = 10^{-3}$ *FACPACK*-PIA requires about 21.7s for the computation of $\tilde{\mathcal{M}}_S$ and $\tilde{\mathcal{M}}_C$. In contrast to this, the dual Borgen plot method even with a smaller control parameter and with a number of 5000 initial tangents takes only about 3s. Hence

control parameters		number of vertices		Comp. times [s] GBP	
rot. angle	d-stopping	\mathcal{M}_S	\mathcal{M}_C	\mathcal{M}_S	\mathcal{M}_C
10^{-1}	10^{-3}	2120	–	3.027	–
10^{-1}	10^{-4}	2102	–	2.958	–
$5 \cdot 10^{-2}$	10^{-3}	3490	–	5.226	–
$5 \cdot 10^{-2}$	10^{-4}	3505	–	5.317	–

Table 4: Computation times of the generalized Borgen plot (GBP) method. The numbers of computed tangents are comparable to the dual Borgen plot method, see Table 2, but much smaller than those of *FACPACK*-PIA, see Table 3. The AFS \mathcal{M}_C cannot be computed with GBP as the magnitude of negative components is too large.

the dual Borgen plot construction is about 7 times faster than the (already fast) *FACPACK*-PIA method. The precision of the results is not affected. The AFS approximations by these methods coincide aside from differences of the magnitude of the boundary precision ε_b .

The high efficiency of the dual Borgen plot method can be explained by the fact that it immediately accesses the structure-forming objects, namely the polygons FIRPOL and INNPOL. In contrast to this, the optimization based *FACPACK*-PIA algorithm explores for each subset of the AFS its boundary by using an adaptive approximation algorithm. This also explains why *FACPACK*-PIA detects a characteristically smaller number of vertices, see Tables 2, 3 and 4. However, the smaller number of vertices does not affect the precision of the boundary approximation which is controlled by the parameter ε_b as the non-detected vertices belong to pairs of edges which are nearly collinear.

4.5.2. Comparison to generalized Borgen plots

Table 4 contains the computation times and number of vertices for the approximate construction of \mathcal{M}_S by the generalized Borgen plot (GBP) method. For the default control parameter settings rot. angle= 0.1 and d-stopping= 10^{-3} the computation of \mathcal{M}_S is completed within 3s. The resulting AFS-sets by GBP are slightly different to the AFS-sets by the dual Borgen plot method due to the different approach to use absolute lower bounds on the acceptable negative matrix entries compared to the relative bounds (10). Further GBP cannot compute the AFS \mathcal{M}_C since the magnitude of negative entries in this experimental data set is too large.

5. Conclusion

MCR methods should be suitable for applications to experimental spectral data including noise, perturbations, baseline drifts and other types of errors and nonlinearities. Data preprocessing steps as baseline corrections, background subtractions or noise-reducing low-rank approximations can introduce small negative entries to the spectral data set. The classical Borgen plots by Borgen and Kowalski cannot handle such data with small negative entries.

The present paper devises a way on how to work with small negative data entries in the dual Borgen plot constructions. The idea is to map the weakened nonnegativity constraints to two modified, namely increased, polygons FIRPOL, which represent in the classical Borgen plot construction exactly the nonnegativity constraints. For given FIRPOL, duality principles help to construct a corresponding modified polygon INNPOL. This reorganized construction of the polygons INNPOL and FIRPOL, which are nesting sets of the AFS, is followed by the dual Borgen plot construction as introduced in the first part of this paper [5]. The resulting overall process combines the elegance of the geometric Borgen plot construction with the numerical robustness of the optimization-based AFS computation methods as triangle enclosure [7], polygon inflation [10, 11] and ray casting [15].

The approach to map the weakened nonnegativity constraint to a modified definition of the polygon FIRPOL makes it possible to implement not only lower bounds on the relative size of negative matrix entries (as used here), but also to use absolute lower bounds as done in the generalized Borgen plot (GBP) method. In general, constraint implementation via a re-definition of the polygons FIRPOL and INNPOL in the Borgen plot constructions appears to be a challenging research topic.

References

- [1] 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.
- [2] 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.
- [3] 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.
- [4] A. Meister. Estimation of component spectra by the principal components method. *Anal. Chim. Acta*, 161:149–161, 1984.
- [5] M. Sawall, A. Jürß, H. Schröder, and K. Neymeyr. Simultaneous construction of dual Borgen plots. I: The case of noise-free data. Technical report, To appear in *J. Chemom.*, DOI: 10.1002/cem.2954., 2017.
- [6] 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.
- [7] A. Golshan, H. Abdollahi, and M. Maeder. Resolution of rotational ambiguity for three-component systems. *Anal. Chem.*, 83(3):836–841, 2011.
- [8] A. Jürß, M. Sawall, and K. Neymeyr. On generalized Borgen plots. II: The line-moving algorithm and its numerical implementation. *J. Chemom.*, 30:636–650, 2016.
- [9] M. Sawall, A. Jürß, H. Schröder, and K. Neymeyr. *On the analysis and computation of the area of feasible solutions for two-, three- and four-component systems*, volume 30 of *Data Handling in Science and Technology*, “Resolving Spectral Mixtures”, Ed. C. Ruckebusch, chapter 5, pages 135–184. Elsevier, Cambridge, 2016.
- [10] 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.
- [11] 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.
- [12] 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.
- [13] 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.
- [14] H. Abdollahi and R. Tauler. Uniqueness and rotation ambiguities in Multivariate Curve Resolution methods. *Chemom. Intell. Lab. Syst.*, 108(2):100–111, 2011.
- [15] M. Sawall and K. Neymeyr. A ray casting method for the computation of the area of feasible solutions for multicomponent systems: Theory, applications and FACPAC-implementation. *Anal. Chim. Acta*, 960:40–52, 2017.
- [16] W.H. Lawton and E.A. Sylvestre. Self modelling curve resolution. *Technometrics*, 13:617–633, 1971.
- [17] H. Schröder, M. Sawall, C. Kubis, D. Jürß, A. Selent, A. Brächer, A. Börner, R. Franke, and K. Neymeyr. Comparative multivariate curve resolution study in the area of feasible solutions. *Chemom. Intell. Lab. Syst.*, 163:55–63, 2017.
- [18] H. Minc. *Nonnegative matrices*. John Wiley & Sons, New York, 1988.
- [19] R.C. Henry. Duality in multivariate receptor models. *Chemom. Intell. Lab. Syst.*, 77(1-2):59–63, 2005.
- [20] R. Rajkó. Natural duality in minimal constrained self modeling curve resolution. *J. Chemom.*, 20(3-4):164–169, 2006.
- [21] 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.
- [22] 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.
- [23] 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.
- [24] H. Abdollahi, M. Maeder, and R. Tauler. Calculation and Meaning of Feasible Band Boundaries in Multivariate Curve Resolution of a Two-Component System. *Anal. Chem.*, 81(6):2115–2122, 2009.
- [25] 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.
- [26] R. Rajkó, H. Abdollahi, S. Beyramysoltan, and N. Omidikia. Definition and detection of data-based uniqueness in evaluating bilinear (two-way) chemical measurements. *Anal. Chim. Acta*, 855:21–33, 2015.
- [27] 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.
- [28] M. Maeder and Y.M. Neuhold. *Practical data analysis in chemistry*. Elsevier, Amsterdam, 2007.
- [29] 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.
- [30] 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.
- [31] M. Sawall, A. Jürß, and K. Neymeyr. FACPAC: A software for the computation of multi-component factorizations and the area of feasible solutions, Revision 1.3. FACPAC homepage: <http://www.math.uni-rostock.de/facpack/>, 2015.