On Generalized Borgen Plots.
I: From convex to affine combinations
and applications to spectral data.

To the memory of Odd S. Borgen (1929–1994).

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

[a]*Universität Rostock, Institut für Mathematik, Ulmenstraße 69, 18057 Rostock, Germany,*
[b]*Leibniz-Institut für Katalyse e.V. an der Universität Rostock, Albert-Einstein-Straße 29a, 18059 Rostock.*

Abstract

In 1985 Borgen and Kowalski [DOI:10.1016/S0003-2670(00)84361-5] published their landmark paper on the geometric construction of feasible regions for nonnegative factorizations of spectral data matrices for three-component systems. These geometric constructions are called Borgen plots. Borgen plots are principally restricted to nonnegative data and are sometimes considered as analytical tool. Major contributions to this theory have been given by Rajkó. In contrast to these geometric constructions, numerical methods to compute the so-called Area of Feasible Solutions (AFS) have been studied by Golshan et al. [DOI: 10.1021/ac102429q] and by Sawall et al. [DOI: 10.1002/cem.2498]. These numerical methods can even treat spectral data which include slightly negative components.

In this work the concept of *Generalized Borgen Plots* is introduced for spectral data which are polluted by small negative entries. The analysis is not restricted to three-component systems, but can be applied to general *s*-component systems. Generalized Borgen plots are identical to the classical Borgen plots for nonnegative data. The analysis in this work also bridges the gap between the different scalings (Borgen norms) used for AFS computations.

The algorithmic procedure of generalized Borgen plots for three-component systems and its implementation in the *FAC-PACK* software are described in the second part of this paper.

Key words: factor analysis, pure component decomposition, nonnegative matrix factorization, Borgen plot, tangent algorithm, spectral recovery.

1. Introduction

The extraction of pure component information from spectroscopic measurements on multicomponent chemical reaction systems is an important problem of analytical chemistry. Chemometric methods are valuable tools to determine not only the number of independent components in the reaction system, but also to extract the concentration profiles and spectra of the pure components.

The starting point of such an analysis is the $k \times n$ spectroscopic data matrix D, whose rows contain a number of k spectra and each spectrum contains n absorbance values with respect to a fixed wavelength grid. The matrix form of the Lambert-Beer law $D = CA$ states that D, aside from small nonlinearities and measurement errors, is a product of a concentration matrix $C \in \mathbb{R}^{k \times s}$

and a spectra matrix $A \in \mathbb{R}^{s \times n}$ where s is the number of independent species. The columns of C represent the concentration profiles of the pure components along the time axis and the rows of A contain the pure component spectra. Due to their physical meaning the components of the three matrices D, C and A are nonnegative numbers.

Here we consider the reverse problem, namely to determine for a given spectral data matrix D just the two unknown factors C and A. Unfortunately, the nonnegative factorization of D is not unique in most cases. This fact is known as the *rotational ambiguity* of nonnegative factorizations of D. Multivariate curve resolution methods resolve this ambiguity problem by using soft and hard constraints (e.g. unimodality, closure of concentrations, smoothness of spectra or solutions, kinetic models and so on) which in many cases allow to extract

a single solution, see also [3].

However, an important question is to determine the set of all feasible nonnegative factors C and A so that the product CA constructs the given spectral data matrix D. In 1971 Lawton and Sylvestre [12] solved this problem for a two-component system by presenting the so-called LS-plots. They showed that there is a one-to-one relation between feasible rows of A and certain regions in the plane which represent admissible (mixing or expansion) coefficients with respect to the basis of singular vectors, see also [17]. A comparable relation holds for the feasible columns of C. In 1985 these ideas were extended to three-component systems by Borgen and Kowalski [5].

Together with a normalization condition the set of all nonnegative factorizations of D for a three-component system can be represented by a bounded subset of the plane. This low-dimensional representation of feasible solutions of the factorization problem is called the *area of feasible solutions* (AFS). A typical example of an AFS is shown in Figure 1, see [11]. In 1985 Borgen and Kowalski described two geometric algorithms to construct the AFS, namely the tangent algorithm and the simplex rotation algorithm. In 2005 Rajkó and István [18] introduced methods of computational geometry to draw Borgen plots for three-component systems. A key condition for these algorithms is the nonnegativity of D as well as of C and A. Borgen plots were originally developed for ideal noise-free bilinear data and mainly for principle-based research work.

The nonnegativity of spectral data is generally a necessary prerequisite which is sometimes violated for experimental spectroscopic data. However, preprocessing steps like background subtraction or the elimination of known pure components from the spectral data can result in small negative elements. Further for spectroscopic measurements, perturbations increase the rank of the spectral data matrix and the low rank approximation matrix can contain negative components. Then the classical Borgen plots cannot be constructed. In addition to the geometric constructive approach by Borgen and Kowalski two alternative numerical algorithms have been developed to approximate the Borgen plots by an AFS for three-component systems. First, the triangle enclosure algorithm has been presented in 2011 [8]. Second, the polygon inflation algorithm has been suggested in 2013 [21, 22]. These two numerical algorithms are based on a different scaling compared to the geometric construction in [18] so that the result can look quite different. An important benefit of the triangle enclosure algorithm and the polygon inflation algorithm is that they can even work with spectral data which include slightly negative components. The survey paper [7] contains a comparative study of Borgen plots, numerical AFS approximations and of the techniques [6, 26] to compute ranges of feasible solutions.

This paper focuses on the nonnegativity constraint on factorizations of D. If further soft or even hard constraints on the solution are added, then the resulting constrained AFS will be smaller, see e.g. [19, 25]. In the extreme case of an AFS which consists only of isolated points, namely one for each component, a unique solution has been found. There exists, however, for our geometric construction of the AFS not always a possible translation of the constraint to associated geometric conditions.

1.1. Topics and aims

The topics of this work are as follows:

1. A generalization of the classical Borgen plots is presented, which allows that the geometric constructive Borgen plots can now be applied to experimental spectroscopic data which is contaminated with small negative components. These negative components can be caused by data preprocessing, low rank approximation or background subtraction. We call the resulting AFS approximations for data with small negative components or for rank-perturbed data the *generalized Borgen plots*. If the perturbations or negative components tend to zero, then the generalized Borgen plots tend to the classical Borgen plots.

2. The mathematical theory behind the classical Borgen plots, which is based on *convex linear combinations*, is generalized to *affine linear combinations*. These affine linear combinations empower the geometric constructive approach to work successfully with spectral data which includes negative entries. (The idea of considering slightly negative components within the factorization problem has already been used in [6] to handle noisy data.)

3. The gap between the geometric constructive Borgen plots [5, 15, 18] and the numerical AFS algorithms in [8, 21, 22] with their different factor scalings is bridged. An explicit transformation between row sum scaling and first singular vector scaling is derived.

4. The mathematical theory is presented for general s-component systems; there is no restriction to $s = 3$.

The algorithmic procedure for the geometric construction of generalized Borgen plots for three component systems is presented in part II of this paper. The new algorithms have been implemented in the *FAC-PACK* software.

1.2. Guideline for the reader

This paper reports on new concepts for the AFS and the underlying mathematical analysis. The reader who is mainly interested in concepts can skip all proofs. The topics and aims of the paper are explained in Section 1.1. Throughout the paper verbal explanations precede the mathematical theorems. Various figures and numerical examples (e.g. in Sections 2.5 and 4) accompany the discussion. In any case, a good starting point for the reader might be experimenting with the (generalized) Borgen plots by using the software provided at:

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

1.3. Organization of this paper

In Section 2 two commonly used factor scalings for the AFS construction are discussed. The equivalence of these representations is proved and formula for their mutual transformation are presented. Section 3 contains the core results of this paper. The mathematical fundamentals of the classical Borgen and Kowalski geometric construction is generalized to data which includes small negative components. The theory is presented in general form for s-component systems and provides the mathematical basis for the geometric construction of the AFS. We call these geometric constructions *generalized Borgen plots*. Finally, Section 4 contains generalized Borgen plots for a model problem and for experimental FT-IR spectroscopic data. The new algorithms are demonstrated for various parameter selections and the results are compared with AFS results from independent methods.

1.4. Notation

The following notation is used in the paper. The references apply to the first usage of the symbol.

$D \in \mathbb{R}^{k \times n}$	spectral data matrix, see Sec. 1.
$C \in \mathbb{R}^{k \times s}$	concentration matrix, see Sec. 1.
$A \in \mathbb{R}^{s \times n}$	spectra matrix, see Sec. 1.
$U \Sigma V^T$	singular value decomposition of D, see Sec. 2.3.
$T \in \mathbb{R}^{s \times s}$	transformation matrix, see (9) and (10).

$t \in \mathbb{R}^{s-1}$	low-dimensional representation of spectra by $t = T(1, 2 : s)$, see (10).
$\mathcal{M}$	the AFS, see Def. 2.4.
$\mathcal{M}_{\varepsilon_C, \varepsilon_A}$	generalized AFS for slightly negative data, see Def. 3.8.
$\overline{D} \in \mathbb{R}^{k \times n}$	spectral data matrix w.r.t. RS-scaling, see Sec. 2.2
$\overline{C} \in \mathbb{R}^{k \times s}$	concentration matrix w.r.t. RS-scaling, see Sec. 2.2.
$\overline{A} \in \mathbb{R}^{s \times n}$	spectra matrix w.r.t. RS-scaling, see Sec. 2.2.
$\overline{U} \, \overline{\Sigma} \, \overline{V}$	singular value decomposition of $\overline{D}$, see Sec. 2.2.
$\overline{T} \in \mathbb{R}^{s \times s}$	transformation matrix w.r.t. RS-scaling, see Eq. (6).
$\overline{t} \in \mathbb{R}^{s-1}$	low-dimensional representation of spectra w.r.t. RS-scaling with $\overline{t} = \overline{T}(1, 2 : s)$, see Sec. 2.2.
$\overline{\mathcal{M}}$	the AFS w.r.t. RS-scaling, see Def. 2.3.
$\overline{\mathcal{M}}_{\varepsilon_C, \varepsilon_A}$	generalized AFS for slightly negative data w.r.t. RS-scaling, see Def. 3.1.
$\Omega \in \mathbb{R}^{k \times k}$	scaling matrix, see Eq. (24).
$\widehat{D} \in \mathbb{R}^{k \times s}$	spectral data matrix w.r.t. FSV-scaling, see Eq. (25).

2. On the representation of feasible factorizations

Let D be a k-by-n spectral data matrix as introduced in Section 1. The number of active and independent species is denoted by s. For the analytical part of the paper the number s is not restricted. For the numerical examples we use $s = 3$. In any case $s \leq \min(k, n)$. All nonnegative factorizations CA of D are to be determined with $C \in \mathbb{R}^{k \times s}$ and $A \in \mathbb{R}^{s \times n}$. If $D = CA$ is a nonnegative factorization, then for any diagonal matrix $\Theta \in \mathbb{R}^{s \times s}$ with strictly positive diagonal elements the factorization

$$D = (C\Theta^{-1})(\Theta A) \qquad (1)$$

is also a nonnegative factorization. Equation (1) means that the ith row or spectrum of A is scaled by the positive factor $\Theta_{i,i}$ and that simultaneously the ith column of C (concentration profile) is scaled by $1/\Theta_{i,i}$. Thus Equation (1) expresses a trivial scaling ambiguity. In Section 2.1 two commonly used standardizations of this scaling are introduced and their implications are analyzed. We still need the definition of the 1-norm.

Definition 2.1. *The 1-norm of a vector $\upsilon \in \mathbb{R}^n$ is the*

sum of the absolute values [10]

$$\|v\|_1 = \sum_{i=1}^{n} |v_i|. \tag{2}$$

2.1. The row sum scaling and the first singular vector scaling

The scaling of the spectral factor A according to (1) has a significant effect on the shape of the AFS, see [15]. Sometimes the scaling is a normalization, e.g. with respect to the 1-norm given by (2) [16]. However, the FSV-scaling, which is introduced below, is just a scaling by a proper positive constant (and is not a normalization in the strict sense of a mathematical norm).

is not related to a normalization. Furthermore, our scope is somewhat more general as we are interested in working with data which include negative elements. Then the 1-norm (2) is substituted by the simple sum of the vector components without taking the absolute values. For nonnegative data the row sum is identical to the 1-norm of just this vector.

The following two scaling variants are commonly used for AFS computations:

1. All rows of A are assumed to be normalized so that $\sum_{j=1}^{n} A(i, j) = 1$ for $i = 1, \ldots, s$. In the following we call this **r**ow **s**um scaling the **RS-scaling**.

2. The rows of A are scaled so that the coefficients of the linear combinations of each row of A with respect to the first right singular vector of D equals 1. For the singular value decomposition [10] of the spectroscopic data matrix D see [14] and Section 2.2. Theorem 2.2 in [22] proves that this coefficient is nonzero under a certain weak assumption on D. In the following we call this **f**irst right **s**ingular **v**ector scaling the **FSV-scaling**.

The RS-scaling is one of the scalings suggested in [5]. Later the RS-scaling is used, e.g., in [18]. This scaling does not require a singular value decomposition of D. In contrast to this, the numerical algorithms to compute the AFS work within the basis of left and right singular vectors of D, see [21]. With respect to this basis the FSV-scaling can be used. In this paper we use the RS-scaling and the FSV-scaling simultaneously. Bars are added to all variables which refer to RS-scaling and bars are skipped for variables which refer to FSV-scaling.

2.2. The AFS for RS-scaling

The term *Borgen plot*, which is coined by Rajkó and István [18], is commonly used for the geometric construction of the AFS which is computed by the algorithm of Borgen and Kowalski [5]. Rajkó and István primarily used a row normalization of A with the 1-norm (2) in order to construct the AFS.

The starting point is a row sum scaling of D resulting in $\overline{D} = \Delta D$ where $\Delta \in \mathbb{R}^{k \times k}$ is the nonnegative diagonal matrix with

$$\Delta(i, i) = \|D(i, :)\|_1^{-1} = 1/(\sum_{\ell=1}^{n} D(i, \ell)). \tag{3}$$

Next a nonnegative factorization $\overline{D} = \overline{C}\,\overline{A}$ is wanted. The factorization of $\overline{D}$ instead of D does not restrict the generality of the approach since $D = \Delta^{-1}\overline{C}\,\overline{A}$. If additionally the row sums of $\overline{A}$ are all equal to 1, then the rows of $\overline{D}$ can be interpreted geometrically as shown in the next lemma.

Lemma 2.2. *Let $\overline{D} \in \mathbb{R}^{k \times n}$ be a nonnegative matrix whose row sums are all equal to 1. The factorization $\overline{D} = \overline{C}\,\overline{A}$ with $\overline{C} \in \mathbb{R}^{k \times s}$ and $\overline{A} \in \mathbb{R}^{s \times n}$ is a nonnegative matrix factorization with all row sums of $\overline{A}$ being equal to 1 if and only if each row of $\overline{D}$ is a convex combination of the rows of $\overline{A}$, $\overline{A} \geq 0$ and the rows of $\overline{C} \geq 0$ are the coefficients of the convex combination. Thus all row sums of $\overline{C}$ are equal to 1.*

Proof. Let $\overline{D} = \overline{C}\,\overline{A}$ be a nonnegative matrix factorization. Then the rows $\overline{D}(i, :)$ of $\overline{D}$ are linear combinations of the rows of $\overline{A}$, since

$$\overline{D}(i, :) = \sum_{\ell=1}^{s} \overline{C}(i, \ell)\overline{A}(\ell, :). \tag{4}$$

The given row scaling of $\overline{D}$ and $\overline{A}$ implies that

$$1 = \sum_{j=1}^{n} \overline{D}(i, j) = \sum_{j=1}^{n} \sum_{\ell=1}^{s} \overline{C}(i, \ell)\overline{A}(\ell, j)$$
$$= \sum_{\ell=1}^{s} \overline{C}(i, \ell) \underbrace{\sum_{j=1}^{n} \overline{A}(\ell, j)}_{=1} = \sum_{\ell=1}^{s} \overline{C}(i, \ell). \tag{5}$$

Thus all row sums of $\overline{C}$ are equal to 1. This proves together with the nonnegativity of the $\overline{C}(i, \ell)$ that Equation (4) is not only a linear combination but even a convex combination of the rows of $\overline{A}$. For the reverse direction with $\overline{C}, \overline{A} \geq 0$, Equation (5) proves the row sum condition for $\overline{A}$ and the convex combinations (4) prove that $\overline{D} = \overline{C}\,\overline{A}$ is a nonnegative matrix factorization. $\square$

The geometric construction of the AFS in the form of Borgen plots and its generalization are explained in Section 3. Next a mathematical description of the AFS with respect to RS-scaling is given. Therefore let a nonnegative matrix $\overline{D} \in \mathbb{R}^{k \times n}$ of the rank s be given. The truncated rank-s singular value decomposition (SVD) reads $\overline{D} = \overline{U}\,\overline{\Sigma}\,\overline{V}^T$ with $\overline{\Sigma} \in \mathbb{R}^{s \times s}$ and orthogonal $\overline{U} \in \mathbb{R}^{k \times s}$ and $\overline{V} \in \mathbb{R}^{s \times n}$; see [10]. The key idea for the mathematical representation of the rotational ambiguity is to insert a regular matrix $\overline{T} \in \mathbb{R}^{s \times s}$ and its inverse in the truncated (SVD)

$$\overline{D} = \overline{U}\,\overline{\Sigma}\,\overline{V}^T = \underbrace{\overline{U}\,\overline{\Sigma}\,\overline{T}^{-1}}_{=:\overline{C}}\ \underbrace{\overline{T}\,\overline{V}^T}_{=:\overline{A}}. \qquad (6)$$

The representation $\overline{A} = \overline{T}\,\overline{V}^T$ shows that the rows of $\overline{T}$ are low-dimensional representations of the possible solutions. Similarly to the identity matrix $\overline{T}^{-1}\overline{T}$ a permutation matrix and its inverse (transposed matrix) can be inserted in the truncated SVD. This shows that the set of all feasible spectra is completely determined by the set of all possible first rows to $\overline{T}$ [21]. Additionally, these s degrees of freedom in the first row of $\overline{T}$ can be further reduced by one degree of freedom by the row sum scaling condition for $\overline{A}$ [5, 18]. It holds that

$$1 = \sum_{i=1}^{n} \overline{A}(1,i) = \sum_{j=1}^{s}\sum_{i=1}^{n} \overline{T}(1,j)\overline{V}(i,j).$$

Thus

$$\overline{T}(1,1) = \frac{1 - \sum_{j=2}^{s}\sum_{i=1}^{n} \overline{T}(1,j)\overline{V}(i,j)}{\sum_{i=1}^{n} \overline{V}(i,1)} \qquad (7)$$

shows that $\overline{T}(1,1)$ is uniquely determined by $\overline{T}(1, 2 : s)$ and $\overline{V}$. Hence the AFS is fully determined by the set of all feasible row vectors $\bar{t} := \overline{T}(1, 2 : s)$. All this results in the following definition.

Definition 2.3. *Let $\overline{D}$ be a rank-s matrix whose row sums are all equal to 1. Let $\overline{D} = \overline{U}\,\overline{\Sigma}\,\overline{V}^T$ be the SVD of $\overline{D}$. The AFS with respect to RS-scaling is the set of row vectors $\bar{t} \in \mathbb{R}^{1 \times s-1}$, for which a regular matrix $\overline{T} \in \mathbb{R}^{s \times s}$ exists with $\overline{T}(1, 2 : s) = \bar{t}$ so that $\overline{C} = \overline{U}\,\overline{\Sigma}\,\overline{T}^{-1}$ and $\overline{A} = \overline{T}\,\overline{V}^T$ are nonnegative matrices. Thus the AFS with respect to RS-scaling, equivalently the Borgen plot, is the set*

$$\overline{\mathcal{M}} := \Big\{ \bar{t} \in \mathbb{R}^{1 \times s-1} : exists\ regular\ \overline{T},\ \overline{T}(1, 2 : s) = \bar{t}, \\ \overline{U}\,\overline{\Sigma}\,\overline{T}^{-1} \ge 0, \overline{T}\,\overline{V}^T \ge 0, \|\overline{T}(i,:)\overline{V}^T\|_1 = 1\ \forall i \Big\}. \qquad (8)$$

2.3. The AFS for FSV-scaling

The genuinely numerical procedures to compute the AFS in [1, 2, 8, 9, 21] do not require an initial row scaling for D prior to the computation of the AFS. Instead the starting point is a truncated rank-s singular value decomposition of $D = U\Sigma V^T$ with $U \in \mathbb{R}^{k \times s}$, $\Sigma \in \mathbb{R}^{s \times s}$ and $V \in \mathbb{R}^{n \times s}$. It is important to note, that the SVD of the row scaled matrix $\overline{D}$ is very different from the SVD of D; there is no simple or even linear transformation between the singular values or singular vectors of these two SVDs. Hence $\overline{U}$ and U are very different. The same holds for $\overline{\Sigma}$, Σ as well as for $\overline{V}$, V.

Similarly, a regular matrix $T \in \mathbb{R}^{s \times s}$ is used in order to define the factors

$$C = U\Sigma T^{-1} \quad \text{and} \quad A = TV^T. \qquad (9)$$

Under some weak assumption on D, namely the irreducibility of $D^T D$, the matrix T can be restricted to have the all-ones vector in its first column, see [21]. In other words the coefficient of the first (normalized) right singular vector equals 1 for each spectrum or row of A. Thus T has the form

$$T = \left(\begin{array}{c|ccc} 1 & t_1 & \cdots & t_{s-1} \\ \hline 1 & & & \\ \vdots & & S & \\ 1 & & & \end{array} \right), \qquad (10)$$

where $t = (t_1, \ldots, t_{s-1}) \in \mathbb{R}^{1 \times s-1}$ is a row vector and with $S \in \mathbb{R}^{(s-1) \times (s-1)}$. The counterpart of Definition 2.3 for FSV-scaling is:

Definition 2.4. *Let $D \in \mathbb{R}^{k \times n}$ be a rank-s matrix with irreducible $D^T D$ let $D = U\Sigma V^T$ be the SVD of D. The AFS $\mathcal{M}$ with respect to FSV-scaling is the set of all row vectors $t \in \mathbb{R}^{1 \times s-1}$ in (10) with regular T so that $C = U\Sigma T^{-1}$ and $A = TV^T$ are nonnegative matrices. The set $\mathcal{M}$ reads*

$$\mathcal{M} = \Big\{ t \in \mathbb{R}^{1 \times s-1} : exists\ regular\ T,\ T(1,:) = (1,t), \\ U\Sigma T^{-1} \ge 0, TV^T \ge 0 \Big\}.$$

2.4. Relation between the AFS for RS-scaling and the AFS for FSV-scaling

The Borgen plot, that is the AFS with respect to RS-scaling, looks different from the Borgen plot with respect to FSV-scaling [16]. This fact makes it difficult to compare the results of the geometric constructive Borgen plots with numerically computed AFS. To close the gap, the following theorem provides a point-by-point nonlinear transformation between these two AFS representations.

5

Theorem 2.5. *For a rank-s matrix $D \in \mathbb{R}^{k \times n}$ with irreducible $D^T D$ let $\overline{\mathcal{M}}$ be the AFS with respect to RS-scaling and $\mathcal{M}$ the AFS with respect to FSV-scaling. Then there is a one-one mapping between $\bar{t} \in \overline{\mathcal{M}}$ and $t \in \mathcal{M}$ so that these points represent the same spectrum or row of A.*

The explicit form of this one-one mapping is as follows. For the given $t, \bar{t} \in \mathbb{R}^{1 \times s-1}$ let

$$x := (1, t),$$

$$\bar{x} := (\gamma, \bar{t}) \quad with \quad \gamma = \frac{1 - \sum_{j=2}^{s} \sum_{i=1}^{n} \bar{t}_{j-1} V(i,j)}{\sum_{i=1}^{n} V(i,1)}. \quad (11)$$

Then the relations

$$\bar{x} = \frac{x V^T \overline{V}}{\|x V^T\|_1}, \quad (12)$$

$$x = \frac{\bar{x} \overline{V}^T V}{\bar{x} \overline{V}^T V(:,1)} \quad (13)$$

allow to transform the AFS $\mathcal{M}$ to $\overline{\mathcal{M}}$ and vice versa.

Proof. According to the definitions of $\mathcal{M}$ and $\overline{\mathcal{M}}$, the vectors t and $\bar{t}$ are associated with matrices T and $\overline{T}$ and the first rows of these matrices have the form (11) with γ from (7). Due to the Perron-Frobenius theory [13] $V(:,1)$ is a sign-constant and nonzero vector [21]. Hence the denominator of γ is not equal to zero. Thus γ exists.

For given T the associated spectra matrix is $A = T V^T \geq 0$ and C is also nonnegative with $D = CA$. Together with Δ from (3) this results in a factorization of $\overline{D}$.

$$\overline{D} = \Delta D = \underbrace{\Delta C}_{=:\widetilde{C}} A. \quad (14)$$

Nonnegativity of Δ implies $\widetilde{C} \geq 0$. The row sum scaling for A can be achieved by the diagonal matrix $R \in \mathbb{R}^{s \times s}$ with the diagonal elements $R(i,i) = 1/\|T(i,:)V^T\|_1$. Hence

$$\overline{D} = \underbrace{\Delta C R^{-1}}_{\overline{C}} \underbrace{RA}_{\overline{A}}$$

is a nonnegative factorization of $\overline{D}$ with respect to RS-scaling. Insertion of $\overline{A} = \overline{T}\,\overline{V}^T$ from (6) and its counterpart (9) in the equality $\overline{A} = RA$ results in

$$\overline{T}\,\overline{V}^T = R T V^T \quad (15)$$

or equivalently in

$$\overline{T} = R T V^T \overline{V}.$$

The first row of this matrix equation is just (12).

To prove the other direction we rewrite (15) as

$$T = R^{-1} \overline{T}\,\overline{V}^T V \quad (16)$$

with the diagonal matrix R. The diagonal matrix R^{-1} serves to scale all components of the first column of T to 1. Hence the diagonal elements $R(i,i)$ are equal to $(\overline{T}\,\overline{V}^T V)_{i,1}$, i.e.

$$R_{i,i} = \overline{T}(i,:)\overline{V}^T V(:,1)$$

and thus $R_{1,1} = \bar{x}\overline{V}^T V(:,1)$. The first row of (16) reads

$$T(1,:) = x = (1/R_{1,1})\overline{T}(1,:)\overline{V}^T V = (1/R_{1,1})\bar{x}\overline{V}^T V.$$

This proves (13). We note that $R_{i,i} = \overline{T}(i,:)\overline{V}^T V(:,1) > 0$ for all $i \in \{1, 2, \ldots, s\}$ if $D^T D$ is irreducible (cf. Theorem 2.2 in [22]). $\square$

2.5. A numerical example

In Figure 1 the AFS with RS-scaling $\overline{\mathcal{M}}$ and the AFS with FSV-scaling $\mathcal{M}$ for an FT-IR data set from the rhodium catalyzed hydroformylation, see [11], are shown. Obviously the AFS segments look quite different. However, they represent the same set of feasible factorizations. For more details on the experimental data set see Section 4. The AFS $\overline{\mathcal{M}}$ has been computed by using geometric algorithms implemented in Matlab and C. The AFS $\mathcal{M}$ has been computed from $\overline{\mathcal{M}}$ with Equation (13).

2.6. Boundedness of the AFS

A necessary prerequisite for the actual computation of the AFS is its boundedness, i.e. its finite extension. In [22] it has been shown that $\mathcal{M}$ is a bounded set if and only if $D^T D$ is an irreducible matrix. In contrast to this the boundedness of $\overline{\mathcal{M}}$ can be proved without any further assumptions.

Lemma 2.6. *The AFS $\overline{\mathcal{M}}$ for an s-component system is a subset of the unit ball of the $\mathbb{R}^{s-1}$ with respect to the Euclidean vector norm $\|\cdot\|_2$. This means that $\|t\|_2 \leq 1$ for all $t \in \overline{\mathcal{M}}$.*

Proof. Let $\bar{t} \in \overline{\mathcal{M}}$. By reason of Equation (8) a matrix $\overline{T} \in \mathbb{R}^{s \times s}$ exists whose first row is $(\gamma, \bar{t})$ and $\overline{A} = \overline{T}\,\overline{V}^T$ is its associated spectra matrix. The row sum normalization for $\overline{A}$ together with the norm inequality $\|x\|_1 \geq \|x\|_2$, see [10], lead to

$$1 = \|\overline{A}(1,:)\|_1^2 \geq \|\overline{A}(1,:)\|_2^2 = \|\overline{T}(1,:)\overline{V}^T\|_2^2 = \|\overline{T}(1,:)\|_2^2$$

$$= \sum_{i=1}^{s} \overline{T}(1,i)^2 \geq \sum_{i=2}^{s} \overline{T}(1,i)^2 = \|\bar{t}\|_2^2,$$

6

AFS $\overline{\mathcal{M}}$ with RS-scaling · **AFS $\mathcal{M}$ with FSV-scaling**

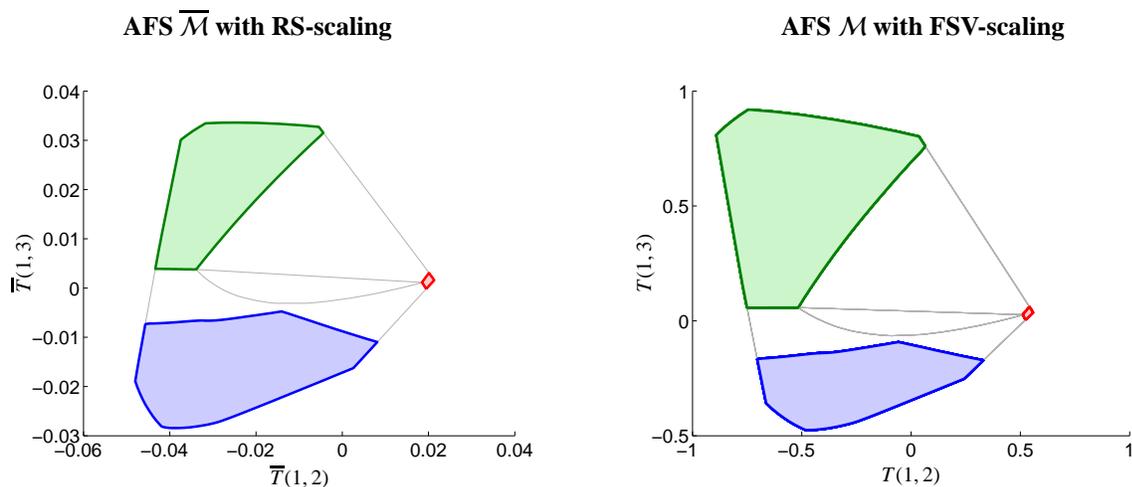Figure 1: AFS $\overline{\mathcal{M}}$ with RS-scaling (left) and $\mathcal{M}$ with FSV-scaling (right) for FT-IR spectroscopic data from rhodium catalyzed hydroformylation process [11]. Each AFS consists of three isolated segments. The gray curves are the polygons FIRPOL and INNPOL.

which proves the proposition. The last equation in the first row follows from the orthogonal invariance of the Euclidean norm. □

Lemma 2.6 proves that a finite AFS with respect to FSV-scaling can be computed irrespectively of whether $D^T D$ is an irreducible matrix. Then the transformation formula of Theorem 2.5 cannot be applied. However, the reducibility of $D^T D$ is nothing which can be expected for practical data aside from the trivial case of completely separated reaction subsystems with separated signal groups in the spectra and concentration profiles.

3. A generalization of Borgen plots

The classical Borgen plots [5] are geometric constructions which can be executed for nonnegative spectroscopic data. The algorithm fails if the data contain negative components which is a disadvantage of the classical Borgen plots. In fact, practical spectral data sets can include small negative elements due to data pre-processing or background subtraction. Further a low rank approximation of a given perturbed spectral data matrix often has negative entries. The truncation of negative elements and their substitution by zero is a possible, but not the best way. Further, measurement errors and noise can be responsible for the non-existence of a nonnegative matrix factorization of D. If small negative entries in C and A can be accepted, then a factorization is more likely to exist.

Here we present a generalization of the Borgen plots which can be applied to spectral data with small negative elements. The meaning and interpretation of the generalized Borgen plots is very similar - however small negative elements in the factors are accepted. We call this generalization of a nonnegative matrix factorization (NMF) an *almost nonnegative matrix factorization*. However, for nonnegative data the generalized algorithm is identical to the classical method of Borgen and Kowalski.

The starting point of the classical Borgen plots for three-component systems is the definition of the two polygons INNPOL and FIRPOL from which the AFS for a given rank-3-matrix can be constructed. FIRPOL is also called the *outer polygon*, see [18]. The following analysis is not restricted to $s = 3$. For $s > 3$ components the two-dimensional polygons INNPOL and FIRPOL turn to be $s - 1$-dimensional polytopes with comparable properties as in the three-component case. Next the classical polygons INNPOL and FIRPOL and the respective higher dimensional polytopes are denoted as INNPOL-RS and FIRPOL-RS in order to express that they refer to the row sum scaling. We also use these polytopes with respect to FSV-scaling. Then these polytopes are called INNPOL-FSV and FIRPOL-FSV.

3.1. The AFS with respect to RS-scaling and for data including negative entries

A spectral data matrix D is considered which may contain small negative matrix elements. These negative elements can origin from a rank-s approximation of the original spectral data or by background subtraction from

7

the measured spectral data. However, the row sums of D must still be positive so that

$$\Delta(i, j) = \begin{cases} 1/(\sum_{\ell=1}^{n} D(i, \ell)) & i = j, \\ 0 & i \neq j \end{cases} \quad (17)$$

is a nonnegative and regular matrix. This definition of Δ uses row sums. A row sum is different from the 1-norm, see (3), of the same row if the row contains negative entries. In the following we work with the row sum scaled matrix $\overline{D} = \Delta D$.

Next the aim is to compute a factorization $\overline{D} = \overline{C}\,\overline{A}$, in which the factors fulfill the componentwise inequalities $\overline{C} \geq -\varepsilon_C$ and $\overline{A} \geq -\varepsilon_A$ for fixed $\varepsilon_C, \varepsilon_A > 0$. This is what we call an approximate NMF. Further, the RS-scaling requires that all row sums of A are equal to 1. These bounds on the factors C and A are similar to those used in [6] and the lower bounds for the relative negative portion of the matrix elements of C and A as used in Section 3.4 of [22]. However, in [22] the maximum norm is used in the denominator whereas here the row sum scaling is used in a comparable fashion.

The AFS $\overline{\mathcal{M}}_{\varepsilon_C, \varepsilon_A}$ for almost nonnegative matrix factorizations is defined as follows:

Definition 3.1. *Let* $\overline{D} \in \mathbb{R}^{k \times n}$ *be a rank-s matrix, whose row sums equal 1. For the truncated rank-s singular value decomposition* $\overline{D} = \overline{U}\,\overline{\Sigma}\,\overline{V}^T$ *and fixed* $\varepsilon_C, \varepsilon_A \geq 0$ *the spectral factor AFS is the set*

$$\overline{\mathcal{M}}_{\varepsilon_C, \varepsilon_A} = \Big\{ \overline{t} \in \mathbb{R}^{1 \times s-1} : \text{exists regular } \overline{T}, \overline{T}(1, 2:s) = \overline{t},$$
$$\overline{U}\,\overline{\Sigma}\,\overline{T}^{-1} \geq -\varepsilon_C, \ \overline{T}\,\overline{V}^T \geq -\varepsilon_A, \quad (18)$$
$$\sum_{j=1}^{s} \sum_{\ell=1}^{n} \overline{T}(i, j)\overline{V}(\ell, j) = 1 \text{ for } i = 1, \ldots, s \Big\}.$$

3.2. Limiting polytopes of the AFS for RS-scaling

For a nonnegative matrix $\overline{D}$ and a nonnegative factorization $\overline{D} = \overline{C}\,\overline{A}$ with all row sums of A being equal to 1, Lemma 2.2 states that the rows of $\overline{D}$ are convex combinations of the rows of $\overline{A}$. The following Lemma 3.3 provides a similar result by using *affine combinations* instead of convex combinations. In words an affine combination is a convex combination in which the sign restriction for the coefficients is omitted.

Definition 3.2. *A vector* $w \in \mathbb{R}^n$ *is an* affine combination *of the vectors* $v_1, v_2, \ldots, v_\ell \in \mathbb{R}^n$ *with* $\alpha_1, \ldots, \alpha_\ell \in \mathbb{R}$*, if*

$$w = \sum_{i=1}^{\ell} \alpha_i v_i \quad and \quad 1 = \sum_{i=1}^{\ell} \alpha_i.$$

The key message of the next lemma is that Lemma 2.2 remains valid for C and A having negative matrix elements.

Lemma 3.3. *Let* $\overline{D} \in \mathbb{R}^{k \times n}$ *be a rank-s matrix, whose row sums equal 1. Then* $\overline{D} = \overline{C}\,\overline{A}$ *is a factorization with* $\overline{C} \in \mathbb{R}^{k \times s}$*,* $\overline{A} \in \mathbb{R}^{s \times n}$ *and componentwise bounds* $\overline{C} \geq -\varepsilon_C$*,* $\overline{A} \geq -\varepsilon_A$ *for* $\varepsilon_C, \varepsilon_A \geq 0$ *with all row sums of* $\overline{A}$ *being equal to 1 if and only if each row of* $\overline{D}$ *is an affine combination of the rows of* $\overline{A}$ *with* $\overline{C} \geq -\varepsilon_C$ *being the matrix of the coefficients and all row sums of* $\overline{C}$ *being equal to 1.*

Proof. The proof follows the lines of the proof of Lemma 2.2. The inequalities $\overline{C}, \overline{A} \geq 0$ are substituted by $\overline{C} \geq -\varepsilon_C$ and $\overline{A} \geq -\varepsilon_A$. In this way convex combinations turn into affine combinations. $\square$

Equation $\overline{A} = \overline{T}\,\overline{V}^T$, see (6), can be interpreted as a low dimensional representation of the rows of $\overline{A}$ by the rows of $\overline{T}$. The key point is that the components of a row of $\overline{T}$ are the coefficients of the linear combinations with respect to the basis of right singular vectors given by the columns of $\overline{V}$. Thanks to Equation (11) the first component of a row of $\overline{T}$ is not important, since $T(i, 1) = \gamma$ can always be recovered from $T(i, 2 : s)$. In other words only the coefficients with respect to the column space of $\overline{V}(:, 2 : s)$ are essential for the analysis. The fundamental relation that the rows of $\overline{D}$ are convex or affine combinations of the rows of $\overline{A} \geq 0$ also holds for the coefficient vectors with respect to the columns of $\overline{V}(:, 2 : s)$. All this gives rise to define two limiting polytopes within this subspace.

Definition 3.4. *The convex hull of the row vectors of* $\overline{D}\,\overline{V}(:, 2 : s)$*, that are the coefficient vectors of the rows of* $\overline{D}$ *with respect to the right singular vectors in* $\overline{V}(:, 2 : s)$*, is called INNPOL-RS.*

INNPOL-RS is a subset of the $\mathbb{R}^{1 \times s-1}$*. For* $s = 3$ *INNPOL-RS is identical to the classical polygon IN-NPOL by Borgen and Kowalski [5] in the case of non-negative* $\overline{D}$ *if all weighting factors* w_k *in [5] are set equal to 1.*

INNPOL is an abbreviation which stands for *inner polygon*. A second polygon (or polytope in the case of higher dimensions) is FIRPOL for *first polygon*. FIRPOL-RS is a larger polytope which includes INNPOL-RS, see Figure 2. FIRPOL-RS encloses the AFS $\overline{\mathcal{M}}_{\varepsilon_C, \varepsilon_A}$ and a part of its boundary is a part of the boundary of the AFS.

The starting point for the definition of FIRPOL-RS is the nonnegativity constraint $A \geq 0$. For the definition of

8

FIRPOL-RS this condition is generalized to $A \geq -\varepsilon_A$ so that for its first row the componentwise inequality

$$\overline{A}(1,:) = \sum_{i=1}^{s} \overline{T}(1,i)(\overline{V}(:,i))^T \geq -\varepsilon_A \qquad (19)$$

holds. The normalization constraint

$$1 = \sum_{i=1}^{n} \overline{A}(1,i) = \sum_{i=1}^{n}\sum_{j=1}^{s} \overline{T}(1,j)\overline{V}(i,j)$$

allows to write $\overline{T}(1,1)$ in the form

$$\overline{T}(1,1) = \frac{1 - \sum_{i=1}^{n}\sum_{j=2}^{s} \overline{T}(1,j)\overline{V}(i,j)}{\sum_{i=1}^{n} \overline{V}(i,1)}. \qquad (20)$$

Inequality (19) is equivalent to

$$\sum_{i=2}^{s} \overline{T}(1,i)(\overline{V}(:,i))^T \geq -\varepsilon_A - \overline{T}(1,1)(\overline{V}(:,1))^T.$$

By substitution of $\overline{T}(1,1)$ with (20) we get

$$\sum_{i=2}^{s} \overline{T}(1,i)(\overline{V}(:,i))^T$$

$$\geq -\varepsilon_A - \left(\frac{1 - \sum_{i=1}^{n}\sum_{j=2}^{s} \overline{T}(1,j)\overline{V}(i,j)}{\sum_{i=1}^{n} \overline{V}(i,1)} \right)(\overline{V}(:,1))^T.$$

This equation can be simplified to

$$\sum_{i=2}^{s} \overline{T}(1,i)\left((\overline{V}(:,i))^T - \frac{\sum_{j=1}^{n} \overline{V}(j,i)(\overline{V}(:,1))^T}{\sum_{j=1}^{n} \overline{V}(j,1)} \right) \qquad (21)$$

$$\geq -\varepsilon_A - \frac{(\overline{V}(:,1))^T}{\sum_{j=1}^{n} \overline{V}(j,1)}.$$

Each component of the vector inequality (21) defines an affine half-space in $\mathbb{R}^{1\times s-1}$ since the first sum on the left side of (21) is an inner product of the vector $\overline{T}(1,2:s)$ with a function depending on $\overline{V}$. (In order to support the understanding of the last argument, the definition of an affine half-space is recapitulated: For a column vector v and a real number α an affine half-space is the set of row vectors t so that $tv \geq \alpha$.) The intersection of all these half-spaces defines the polytope FIRPOL-RS.

Definition 3.5 (Generalized set FIRPOL-RS). *The set*

$$\left\{ \overline{t} \in \mathbb{R}^{1\times s-1} : \sum_{i=2}^{s} \overline{t}_{i-1}\left((\overline{V}(:,i))^T - \frac{\sum_{j=1}^{n} \overline{V}(j,i)(\overline{V}(:,1))^T}{\sum_{j=1}^{n} \overline{V}(j,1)} \right) \right.$$

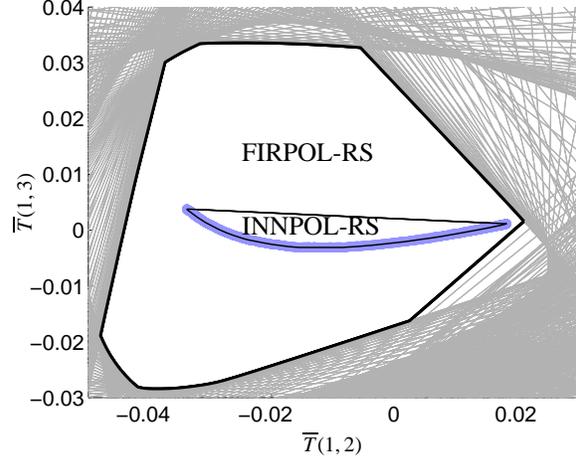$$\left. \geq -\varepsilon_A - \frac{(\overline{V}(:,1))^T}{\sum_{j=1}^{n} \overline{V}(j,1)} \right\}$$

Figure 2: INNPOL-RS and FIRPOL-RS for the AFS $\overline{\mathcal{M}}$ for the FT-IR spectral data, see Section 4, and $\varepsilon_A = -1.1 \cdot \min(\min(D)) = 1.0054 \cdot 10^{-4}$. The gray lines are the boundaries of the half-planes defining FIRPOL-RS. The blue points are the coefficient vectors of the rows of $\overline{D}$ with respect to the right singular vectors $V(:,2)$ and $V(:,3)$.

is called the polytope FIRPOL-RS. Therein $\overline{t}_i$ is the i-th component of the vector $\overline{t}$. For $\varepsilon_A = 0$ and $s = 3$ the polytope FIRPOL-RS is identical to the classical two-dimensional polygon FIRPOL of Borgen and Kowalski for the case that all weighting factors of the standardization in [5] are set equal to 1.

By Lemma 2.6 the polytope FIRPOL-RS is bounded. If $\overline{D}$ has negative matrix elements, then $\overline{D} \geq -\varepsilon_A$ is a necessary condition which guarantees that the geometric constructions work properly; otherwise at least one vertex of INNPOL-RS would be located outside FIRPOL-RS. In Figure 2 INNPOL-RS and FIRPOL-RS are shown for $\varepsilon_A = 1.0054 \cdot 10^{-4}$ and for FT-IR spectroscopic data from [11], see Section 4 for details.

3.3. Geometric characterization of the AFS $\overline{\mathcal{M}}_{\varepsilon_C,\varepsilon_A}$

Theorem 3.6. *The existence of nonnegative matrix factorizations (see case I) and of almost nonnegative matrix factorizations (see case II) for an s-component system is related to a geometric property of s points in FIRPOL-RS.*

I: *Let $\overline{D}$ be a nonnegative matrix whose row sums are all equal to 1. Then $\overline{D} = \overline{C}\overline{A}$ is a nonnegative matrix factorization with RS-scaled factor $\overline{A}$ if and only if for s points $\overline{t}^{(\ell)} \in \mathbb{R}^{1\times s-1}$ in FIRPOL-RS with $\overline{A} = \overline{T}\,\overline{V}^T$ and with $\overline{T}(\ell, 2:s) = \overline{t}^{(\ell)}$ for $\ell = 1,\ldots,s$ the convex hull of these s points $\overline{t}^{(\ell)}$*

9

(being a simplex) includes the polytope INNPOL-RS.

II: *Let $\overline{D}$ be a matrix with negative entries so that $\overline{D} \geq -\varepsilon_A$ for a proper $\varepsilon_A \geq 0$. The row sums of $\overline{D}$ are all equal to 1. Then $\overline{D} = \overline{C}\,\overline{A}$ is a matrix factorization with $\overline{C} \geq -\varepsilon_C$ and $\overline{A} \geq -\varepsilon_A$ and with a RS-scaled factor $\overline{A}$ if and only if for s points $\overline{t}^{(\ell)} \in \mathbb{R}^{1 \times s-1}$ in FIRPOL-RS with $\overline{A} = \overline{T}\,\overline{V}^T$ and $\overline{T}(\ell, 2 : s) = \overline{t}^{(\ell)}$ for $\ell = 1, \ldots, s$ the affine hull of these s points $\overline{t}^{(\ell)}$ includes INNPOL-RS. Therein the affine hull is based on coefficients that are all equal to or greater than $-\varepsilon_C$.*

Proof. Only case II is to be proved. Then the proof of case I follows by setting $\varepsilon_A = \varepsilon_C = 0$.

In order to prove the forward implication the starting point is the matrix factorization

$$\overline{D} = \underbrace{\overline{U}\,\overline{\Sigma}\,\overline{T}^{-1}}_{C}\,\underbrace{\overline{T}\,\overline{V}^T}_{A}$$

with $\overline{C} \geq -\varepsilon_C$ and $\overline{A} \geq -\varepsilon_A$. Each of the s rows of $\overline{T}$ is related to one point in FIRPOL-RS as $\overline{A} \geq -\varepsilon_A$ is satisfied. According to Lemma 3.3 each row of $\overline{D}$ is an affine combination of the rows of $\overline{A}$. The same property holds for the coefficient vectors with respect to the columns of $\overline{V}(:, 2 : s)$, which is shown next. These coefficient vectors of the rows $\overline{D}(i, :)$ are the vectors $\overline{D}(i, :)\overline{V}(:, 2 : s)$. Similarly $\overline{A}(i, :)\overline{V}(:, 2 : s) = \overline{T}(i, 2 : s)$ are these coefficient vectors of the rows of $\overline{A}$. From the affine combination $\overline{D}(i, :) = \sum_{j=1}^{s} \overline{C}(i, j)\overline{A}(j, :)$ right multiplication with $\overline{V}(:, 2 : s)$ yields

$$\overline{D}(i, :)\overline{V}(:, 2 : s) = \sum_{j=1}^{s} \overline{C}(i, j)\overline{A}(j, :)\overline{V}(:, 2 : s). \quad (22)$$

The coefficients $\overline{C}(i, j)$ of this affine combination are still all greater than or equal to $-\varepsilon_C$. Equation (22) shows that representative vectors of the rows of $\overline{D}$ are just affine combinations of the representative vectors of the rows of $\overline{A}$. This guarantees that convex hull INNPOL-RS of the rows of $\overline{D}(i, :)\overline{V}(:, 2 : s)$ is also contained in the affine hull of the rows of $\overline{A}(j, :)\overline{V}(:, 2 : s)$.

For the backward implication we consider the s points $\overline{t}^{(\ell)}$ in $\mathbb{R}^{1 \times s-1}$ with the given properties. Together with the RS-scaling constraint for the factor $\overline{A}$, Equation (11) allows to compute a matrix $\overline{T} \in \mathbb{R}^{s \times s}$ which is unique aside from permutations of its rows and from which $\overline{A} = \overline{T}\,\overline{V}^T$ can be gained. By definition of FIRPOL-RS the inequality $\overline{A} = \overline{T}\,\overline{V}^T \geq -\varepsilon_A$ holds; otherwise at least one point would be located outside of

FIRPOL-RS. The matrix $\overline{T}$ is regular since $\overline{D}$ is a rank-s-matrix and $\overline{D}$ and $\overline{T}\,\overline{V}^T$ span the same column space.

As INNPOL-RS is contained in the affine sums of the $\overline{t}^{(\ell)}$, the rows of $\overline{D}\,\overline{V}(:, 2 : s)$ can be written as affine combinations of the rows of $\overline{A}\,\overline{V}(:, 2 : s)$ with expansion coefficients which form the matrix $\overline{C} \geq -\varepsilon_C$. Analogously, the same holds for the row vectors of $\overline{D}$ and the row vectors of $\overline{C}$. Lemma 3.3 proves that this implies a matrix factorization $\overline{D} = \overline{C}\,\overline{A}$ with the componentwise lower bounds for the matrices $\overline{C}$ and $\overline{A}$. $\square$

Obviously a change of ε_C or ε_A changes the AFS. The following lemma proves the monotonicity principle that an increasing ε_C or ε_A also increases the AFS.

Lemma 3.7. *If $0 < \varepsilon_C \leq \widetilde{\varepsilon_C}$ and $0 < \varepsilon_A \leq \widetilde{\varepsilon_A}$, then the AFS $\overline{\mathcal{M}_{\varepsilon_C, \varepsilon_A}}$ is a subset of the AFS $\overline{\mathcal{M}_{\widetilde{\varepsilon_C}, \widetilde{\varepsilon_A}}}$, i.e.*

$$\overline{\mathcal{M}_{\varepsilon_C, \varepsilon_A}} \subseteq \overline{\mathcal{M}_{\widetilde{\varepsilon_C}, \widetilde{\varepsilon_A}}}.$$

Proof. According to Definition 3.1 a point $\overline{t} \in \overline{\mathcal{M}_{\varepsilon_C, \varepsilon_A}}$ is associated with a matrix $T \in \mathbb{R}^{s \times s}$. This $\overline{T}$ results in the factorization $\overline{D} = \overline{C}\,\overline{A}$ with the element-wise inequalities $\overline{C} \geq -\varepsilon_C$ and $\overline{A} \geq -\varepsilon_A$. Then $\overline{C} \geq -\widetilde{\varepsilon_C}$ and $\overline{A} \geq -\widetilde{\varepsilon_A}$ are also valid. Thus $\overline{t} \in \overline{\mathcal{M}_{\widetilde{\varepsilon_C}, \widetilde{\varepsilon_A}}}$ holds. This proves the proposition. $\square$

3.4. The AFS with respect to FSV-scaling and for data including negative entries

Next the geometric construction principles of the AFS are formulated with respect to FSV-scaling. Our goal is to build a bridge from the geometric-constructive Borgen plots to the numerical algorithms for AFS computations. The latter algorithms typically work within the basis of the left and right singular vectors of D where the FSV-scaling is the natural choice. We show that the geometric construction of the AFS can be adapted to a singular vector basis representation. Additionally a generalization is presented for the AFS construction with respect to FSV-scaling and data including small negative entries.

Let $D = CA$ be a nonnegative factorization and $D = U\Sigma V^T$ be a singular value decomposition (SVD) of the rank-s matrix D with $C, U \in \mathbb{R}^{k \times s}$, $A, V \in \mathbb{R}^{s \times n}$ and $\Sigma \in \mathbb{R}^{s \times s}$. Then $A = TV^T$ and $C = U\Sigma T^{-1}$ for a suitable regular matrix $T \in \mathbb{R}^{s \times s}$. Here we assume that D is a nonnegative matrix and that $D^T D$ is an irreducible matrix. Then the first right singular vector $V(:, 1)$ and the first left singular vector $U(:, 1)$ are sign-constant vectors, i.e. all components are either strictly positive or strictly negative; see [22] for the mathematical background and [13] for the Perron-Frobenius theory. Without loss of generality the SVD can be assumed

to deliver $U(:, 1) > 0$ and $V(:, 1) > 0$; otherwise the two componentwise negative vectors are multiplied by -1 which again results in an SVD.

The SVD of D allows to write the rows of D as linear combinations of the right singular vectors

$$D(i, :) = \sum_{j=1}^{s} w_{i,j} V^T(j, :) \quad (23)$$

with the coefficients $w_{i,j} = (U\Sigma)_{i,j}$. Then the first coefficients $w_{i,1}$ for $i \in \{1, 2, \ldots, s\}$ are positive since with nonzero $D(:, i) \geq 0$ and $V(:, 1) > 0$ one gets

$$0 < D(i, :)V(:, 1) = \Big(\sum_{j=1}^{s} w_{i,j} V^T(j, :)\Big)V(:, 1)$$
$$= \sum_{j=1}^{s} w_{i,j} \delta_{j,1} = w_{i,1}.$$

Therein $\delta_{i,j}$ is the Kronecker delta (which is 0 for $i \neq j$ and 1 for $i = j$). If D is only slightly polluted by negative elements, then a continuity argument shows that $w_{i,1} > 0$ still holds if the perturbation is small enough. The diagonal matrix Ω with the diagonal elements

$$\Omega(i, i) = w_{i,1}^{-1}, \quad i = 1, \ldots, k, \quad (24)$$

or in short form $\Omega = (\mathrm{diag}(W(:, 1)))^{-1}$, is used to define the scaled data matrix

$$\widehat{D} = \Omega D = RV^T. \quad (25)$$

Therein the matrix $R \in \mathbb{R}^{k \times s}$ satisfies that $R(i, 1) = 1$ for all $i \in \{1, 2 \ldots, k\}$.

Definition 3.8. *Let $D \in \mathbb{R}^{k \times n}$ be a rank-s matrix with the truncated rank-s singular value decomposition $D = U\Sigma V^T$. Further let $\varepsilon_C, \varepsilon_A \geq 0$ be given so that the matrix $\Omega \in \mathbb{R}^{k \times k}$ in (24) has positive diagonal elements. Then the spectral factor AFS for FSV-scaling is the set*

$$\mathcal{M}_{\varepsilon_C, \varepsilon_A} = \Big\{ t \in \mathbb{R}^{1 \times s-1} : \ exists \ regular \ T, \ T(1, :) = (1, t),$$
$$\Omega U\Sigma T^{-1} \geq -\varepsilon_C, TV^T \geq -\varepsilon_A \Big\}.$$

3.5. Limiting polytopes of the AFS for FSV-scaling

The counterpart of Definition 3.4 for FSV-scaling reads as follows.

Definition 3.9. *The convex hull of the row vectors of $\Omega D V(:, 2 : s)$ is called the polytope INNPOL-FSV. (These row vectors are the vectors of expansion coefficients of the rows of $\widehat{D} = \Omega D$ by (25) with respect to the columns of $V(:, 2 : s)$.)*

In order to define the polytope FIRPOL-FSV we consider the constraint $A = TV^T \geq -\varepsilon_A$ on acceptable negative entries of A. This inequality together with the scaling condition $T(:, 1) = e^{(s)}$ yields for $i = 1, \ldots, s$

$$\sum_{j=2}^{s} T(i, j)(V(:, j))^T \geq -V(:, 1)^T - \varepsilon_A e^{(s)}. \quad (26)$$

Therein $e^{(s)} = (1, \ldots, 1)$ is the all-ones vector in $\mathbb{R}^s$. As explained after Lemma 2.2 only all possible first rows of T for $i = 1$ must be recorded in order to construct the AFS. The resulting set of feasible first rows is identical to the set of all possible rows of T. This fact can be proved by means of a permutation argument. The set of row vectors $t = T(1, 2 : s) \in \mathbb{R}^{1 \times s-1}$ which satisfy (26) defines the polytope FIRPOL-FSV.

Definition 3.10. *Let $D \geq -\varepsilon_A$ and let $D^T D$ be an irreducible matrix. Then the polytope*

$$\{t \in \mathbb{R}^{1 \times s-1} : \sum_{i=2}^{s} t_{i-1}(V(:, i))^T \geq -(V(:, 1))^T - \varepsilon_A e^{(s)}\}$$

is an intersection of k half-spaces and is called FIRPOL-FSV. Therein $t_1, \ldots, t_{s-1}$ are the components of $t \in \mathbb{R}^{1 \times s-1}$.

A proof for the boundedness of FIRPOL-FSV is given in [22]; this proof uses the irreducibility of $D^T D$. If at least one inequality of the matrix inequality $D \geq -\varepsilon_A$ is violated, then at least one vertex of INNPOL-FSV is outside of FIRPOL-FSV and the geometric algorithm for the AFS construction cannot work.

3.6. Geometric characterization of the AFS $\mathcal{M}_{\varepsilon_C, \varepsilon_A}$

Theorem 3.6 shows how the AFS $\overline{\mathcal{M}}_{\varepsilon_C, \varepsilon_A}$ can be geometrically constructed by means of the polytopes INNPOL-RS and FIRPOL-RS. The following theorem shows that analogous relations hold for the geometric construction of $\mathcal{M}_{\varepsilon_C, \varepsilon_A}$ with the polytopes INNPOL-FSV and FIRPOL-FSV.

Theorem 3.11. *The existence of a nonnegative matrix factorization (see case I) or of a matrix factorization which can include small negative components (see case II) for an s-component system is related to a geometric property of s point in FIRPOL-FSV.*

I: Let D be a nonnegative matrix. Then $D = CA$ is a nonnegative matrix factorization with an FSV-scaled factor A if and only if for s points $t^{(\ell)} \in \mathbb{R}^{1 \times s-1}$ in FIRPOL-FSV with $A = TV^T$ with $T(\ell, 2 : s) = t^{(\ell)}$ for $\ell = 1, \ldots, s$ the convex hull of these s points (being a simplex) includes the polytope INNPOL-FSV.

11

II: Let D be a matrix with negative entries so that $D \geq -\varepsilon_A$ for a proper $\varepsilon_A \geq 0$. Let the diagonal matrix Ω be given according to (24). Then $D = CA$ is a matrix factorization with $\Omega C \geq -\varepsilon_C$ and $A \geq -\varepsilon_A$ and with a FSV-scaled factor A if and only if for s points $t^{(\ell)} \in \mathbb{R}^{1 \times s-1}$ in FIRPOL-FSV with $A = TV^T$ and $T(\ell, 2 : s) = t^{(\ell)}$ for $\ell = 1, \ldots, s$ the affine hull of these s points $t^{(\ell)}$ includes INNPOL-FSV. Therein the affine hull is based on coefficients that are all equal to or greater than $-\varepsilon_C$.

Proof. Only the second case is to be proved. From this the proof for the case I follows by setting $\varepsilon_A = \varepsilon_C = 0$.

For the forward implication the starting point is a factorization

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

with $A \geq -\varepsilon_A$ and $\Omega C \geq -\varepsilon_C$ according to Definition 3.8 so that the first row of T has the form $T(1, :) = (1, t)$ for the given t. This T is the source for the s row vectors $t^{(\ell)} = T(\ell, 2 : s)$, $\ell = 1, \ldots, s$. All these $t^{(\ell)}$ are elements of FIRPOL-FSV as $A \geq -\varepsilon_A$ is satisfied. In order to show the affine hull condition, the ith row of the identity $\widehat{D} = \Omega CA$ from (25) is written in the form

$$\widehat{D}(i, :) = \Omega(i, :)CA = \sum_{j=1}^{s} \underbrace{\Omega(i, i)C(i, j)}_{=:\widehat{C}(i,j)} A(j, :). \quad (27)$$

This is a linear combination of the ith row of $\widehat{D}$ by the rows of A. The sum with respect to j of the coefficients $\widehat{C}(i, j)$ is equal to 1 which is shown next. Equation (25) guarantees that $1 = \widehat{D}(i, :)V(:, 1)$ so that

$$1 = \widehat{D}(i, :)V(:, 1) = \widehat{C}(i, :)AV(:, 1) = \widehat{C}(i, :)TV^T V(:, 1)$$

$$= \widehat{C}(i, :)Te_1 = \widehat{C}(i, :)(1, \ldots, 1)^T = \sum_{j=1}^{s} \widehat{C}(i, j).$$

This proves together with (27) that the rows of $\widehat{D}$ are affine combinations of the rows of A. An analogous relation holds for the vectors of expansion coefficients of the rows of $\widehat{D}$ with respect to the singular vectors $V(:, 2), \ldots, V(:, s)$. By right multiplication of (27) with $V(:, 2 : s)$ we obtain

$$\widehat{D}(i, :)V(:, 2 : s) = \sum_{j=1}^{s} \widehat{C}(i, j)A(j, :)V(:, 2 : s). \quad (28)$$

According to Definition 3.9 the vectors $\widehat{D}(i, :)V(:, 2 : s)$ are just the vertices of the convex polytope INNPOL-FSV. Further, for the vectors $A(j, :)V(:, 2 : s)$ it holds

that

$$A(j, :)V(:, 2 : s) = T(j, :)V^T V(:, 2 : s) = T(j, 2 : s)$$

which are the AFS representatives of the rows of A. Thus (28) shows that each vertex of INNPOL-FSV can be represented as an affine combination of the $T(j, 2 : s)$ with coefficients $(\Omega D)_{i,j} \geq -\varepsilon_C$.

In order to prove the reverse direction the s row vectors $t^{(\ell)} \in \mathbb{R}^{1 \times s-1}$ are written in the rows $1, \ldots, s$ of $T \in \mathbb{R}^{s \times s}$. The first column of T is the all-ones vector. Then $A = TV^T \geq -\varepsilon_A$ holds since all the s vectors are from FIRPOL-FSV. Each point of INNPOL-FSV is a feasible affine combination of the given points and hence this is also true for the rows of $\widehat{D} = WD$, see Equation (27). The associated coefficients of the affine combinations are the elements of the matrix $\widehat{C} = \Omega C$. This completes the proof. $\qquad \square$

Lemma 3.7 states how a variation of ε_C and ε_A changes the AFS $\overline{\mathcal{M}}_{\varepsilon_C, \varepsilon_A}$. An analogous statement is true for $\mathcal{M}_{\varepsilon_C, \varepsilon_A}$.

Lemma 3.12. *If $\varepsilon_C \leq \widehat{\varepsilon_C}$ and $\varepsilon_A \leq \widehat{\varepsilon_A}$, then the AFS $\mathcal{M}_{\varepsilon_C, \varepsilon_A}$ is a subset of the AFS $\mathcal{M}_{\widehat{\varepsilon_C}, \widehat{\varepsilon_A}}$, i.e.*

$$\mathcal{M}_{\varepsilon_C, \varepsilon_A} \subseteq \mathcal{M}_{\widehat{\varepsilon_C}, \widehat{\varepsilon_A}}.$$

The proof is more or less a simple transcription of the proof of Lemma 3.7.

4. Numerical Results

The mathematical fundamentals of generalized Borgen plots in Sections 2 and 3 are now supplemented by a study of sample problems. The practical implementation of these geometric constructions in the form a tangent- and line-moving-algorithm are the topic of part II of this paper. Next a model problem with variable amounts of noise, see Section 4.1, and experimental FT-IR data set from the hydroformylation process, see Section 4.2, are investigated. Different methods are applied to compute the AFS and the results are compared.

4.1. Data set I: A model problem

We consider the concentration profiles

$$c_1(t) = \exp(-(t - 20)^2/150)$$
$$c_2(t) = \exp(-(t - 50)^2/200)$$
$$c_3(t) = \exp(-(t - 70)^2/250)$$

12

and pure component spectra

$$a_1(x) = e^{-(x-50)^2/500} + 0.5e^{-(x-125)^2/500} + 0.1$$

$$a_2(x) = e^{-(x-100)^2/750} + 0.4e^{-(x-100)^2/1500} + 0.15$$

$$a_3(x) = e^{-(x-150)^2/1000} + 0.3e^{-(x-75)^2/2500} + 0.2$$

with $0 \le t \le 100$ and $0 \le x \le 200$. The discretization parameters along the time axis and along the frequency axis are each equal to 0.5. This results in a 201×401 spectral data matrix D_0. Figure 3 shows the $c_i(t)$ and $a_i(t)$ for $i = 1, \ldots, 3$.

Standard normal distributed noise σ is added to D_0

$$D(i, j) := \max\left((1 + \sigma)D_0(i, j), 0\right) \qquad (29)$$

with noise levels $\sigma \in \{0, 0.05, 0.15\}$. The spectral and the concentrational AFS, i.e. the AFS for the factor A and the factor C, are computed with the polygon inflation algorithm from the *FAC-PACK* software. These AFS approximations are compared with the classical Borgen plot and with the generalized Borgen plots, see Figures 4 and 5. The three pure components are marked in all these figures by small black circles. The noise model (29) guarantees that the spectral data matrices D are always nonnegative. However, the rank of these matrices is in general larger than 3 if $\sigma > 0$. The classical Borgen plots can still be constructed for these matrices, but for $\sigma > 0$ the pure components are partly not included in the AFS segment, see e.g. the concentrational AFS in Figure 5 for $\sigma = 0.15$. Hence the "true" solution is not contained within the AFS. This is a serious deficiency of the classical Borgen plots. The Hausdorff distances of the different AFS sets are listed in Table 1. The AFS by the polygon inflation algorithm of *FAC-PACK* is not completely identical to the generalized Borgen plot due to the different approaches to bound the acceptable negative entries in C and A. On the one hand, *FAC-PACK* uses a bound for row-wise relative portion of negative entries while, on the other hand, the generalized Borgen plot uses the absolute bounds $C \ge -\varepsilon_C$ and $A \ge -\varepsilon_A$. This explains why the AFS generated by the polygon inflation algorithm is slightly different from the generalized Borgen plot. Therefore the Hausdorff distances should be interpreted with caution. In any case the generalized Borgen plots in *Figures 4 and 5 show the correct positions of the pure components.*

4.2. Data set II: FT-IR data from hydroformylation

An experimental spectral data set is considered from the rhodium catalyzed hydroformylation process; see [11] for experimental details. This data set comes with a number of $k = 1045$ FT-IR spectra each with $n = 664$ spectral channels. The reaction subsystem consists of three independent components, namely the olefin, the acyl complex and the hydrido complex.

In order to illustrate the influence of the parameters ε_C and ε_A the AFS for the spectral factor and for RS-scaling is shown in Figure 6 for six parameter settings. We explicitly emphasize that the parameters ε_C and ε_A cannot be set arbitrarily. However, these test calculations demonstrate how increasing or decreasing ε_C or ε_A changes the form of the AFS. In Figure 6 the three black circles in each AFS represent the positions of the final pure component spectra which have been determined by means of a kinetic model in [11]. The smallest element of $\overline{D}$ is -0.000101. Negative matrix elements are caused by data preprocessing. For this problem, a background spectrum of the solvent n-hexane has been subtracted and a rank-3 approximation of the spectral data matrix has been computed. Due to this smallest negative matrix element of $\overline{D}$ the control parameter ε_A has to be equal or greater than $-\min(\overline{D}) = 0.000101$. Otherwise INNPOL-RS would have at least one vertex outside of FIRPOL-RS. Then Theorem 3.6 would not allow even a single matrix factorization $\overline{D} = \overline{C}\,\overline{A}$ which fulfills the given lower bounds for the matrix factors $\overline{C}$ and $\overline{A}$. In Figure 2 the top red arrow indicates that decreasing ε_A is prohibited. Further for all AFS sets in Figure 2 the polygons INNPOL-RS and FIRPOL-RS are also drawn. The outer polygon FIRPOL-RS increases if ε_A increases; this can best be seen in the right lower part of the AFS. For increasing ε_C the AFS grows as affine combinations with smaller affine coefficients, that are the matrix elements of $\overline{C}$), are acceptable. The polygon INNPOL-RS remains always unchanged as its definition does not depend on ε_C and ε_A. See also Lemma 3.7 and Lemma 3.12 for the growth of the AFS with increasing ε_A and ε_C.

The computation times for the tangent- and line-moving-algorithms are listed in Table 2. We have used a standard PC with a 2.93GHz Intel CPU and with 8 GB RAM. The program code is written in C and uses the Matlab graphical user interface of the *FAC-PACK* software. For $\varepsilon_C = 0$ and varying ε_A there is only a minor variation of the computation times. With increasing ε_A the area of FIRPOL increases, but this does not complicate the tangent- or line-moving-algorithm. If $\varepsilon_C > 0$ the line-moving-algorithm is used. Then the computation time increases with a growing area of FIRPOL.

4.3. FAC-PACK software for AFS computation

The geometric algorithms for the construction of Borgen plots are a part of the revision 1.2 (appeared in
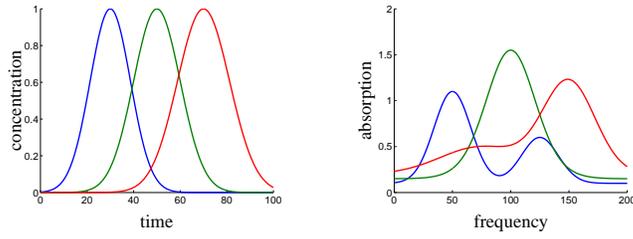
Figure 3: The model problem. Left: Concentration profiles. Right: Pure component spectra. (1,2,3)=(Blue, Green, Red).
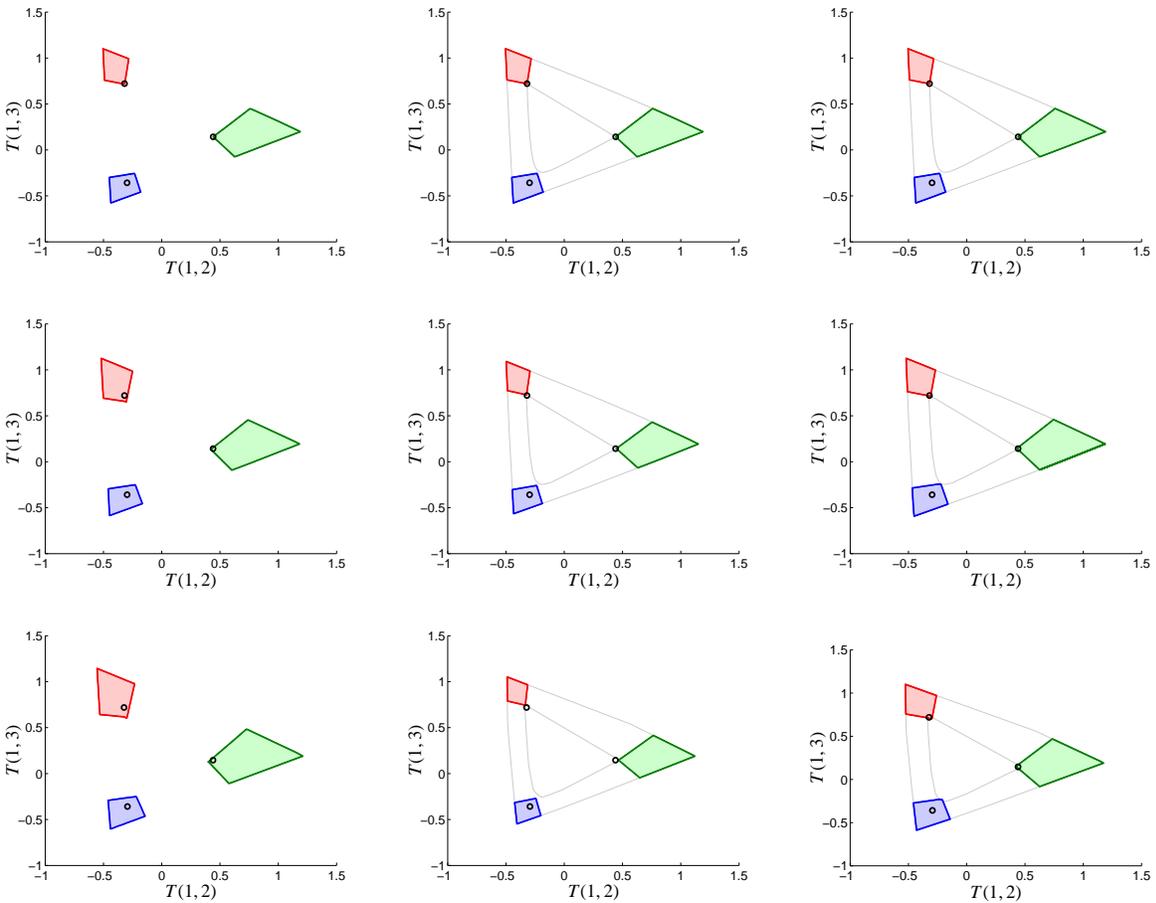

Figure 4: The spectral AFS of the model problem for three noise levels. First row: $\sigma = 0$, second row: $\sigma = 0.05$, third row: $\sigma = 0.15$. First column: The spectral AFS computed by the polygon inflation algorithm of the *FAC-PACK* software, second column: classical Borgen plot with $\varepsilon_C = \varepsilon_A = 0$, third column: generalized Borgen plot with ε_C and ε_A as listed in Table 1. In each AFS the three pure components are marked by small black circles. For $\sigma > 0$ not all pure components are located within the AFS segments of the classical Borgen plots.
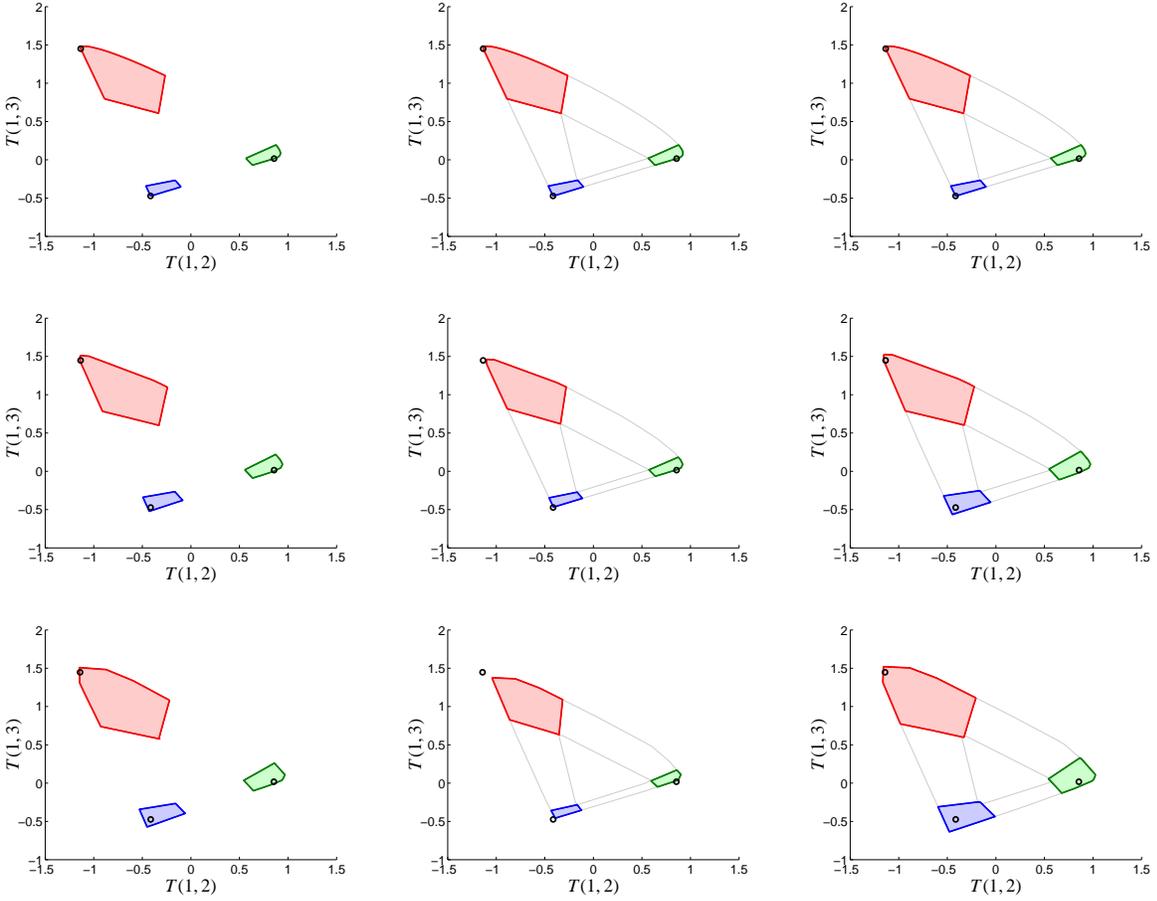
14

Figure 5: The concentrational AFS of the model problem for three noise levels. First row: $\sigma = 0$, second row: $\sigma = 0.05$, third row: $\sigma = 0.15$. First column: The spectral AFS computed by the polygon inflation algorithm of the *FAC-PACK* software, second column: classical Borgen plot with $\varepsilon_C = \varepsilon_A = 0$, third column: generalized Borgen plot with ε_C and ε_A as listed in Table 1. In each AFS the three pure components are marked by small black circles. For $\sigma = 0.15$ all three pure components are not located within the AFS segments of the classical Borgen plots.

σ	PI $\varepsilon_{\text{neg. entr.}}$	gBp parameters	Hausdorff PI-gBp	Hausdorff Bp–gBp	Hausdorff PI–Bp
0	0	$\varepsilon_C = 0, \varepsilon_A = 0$	0.0059	0.0	0.0059
0.05	0.01	$\varepsilon_C = 0.01, \varepsilon_A = 0.002$	0.1273	0.1400	0.0833
0.15	0.0275	$\varepsilon_C = 0.01, \varepsilon_A = 0.005$	0.1487	0.1473	0.1532
0	0	$\varepsilon_C = 0, \varepsilon_A = 0$	0.0342	0.0	0.0342
0.05	0.01	$\varepsilon_C = 0.01, \varepsilon_A = 0.005$	0.1816	0.1826	0.0632
0.15	0.0275	$\varepsilon_C = 0.02, \varepsilon_A = 0.012$	0.1729	0.2121	0.1697

Table 1: Hausdorff distances for three noise levels σ for the spectral AFS, i.e. the matrix factor A, in the rows 2–4 and for the concentrational AFS, i.e. the matrix factor C, in the rows 5–7. Column 1: Noise level. Column 2: *FAC-PACK* polygon inflation (PI) parameter $\varepsilon_{\text{neg. entr.}}$. Column 3: Generalized Borgen plot (gBp) control parameters. Column 4: Hausdorff distances of the AFS computed by polygon inflation (PI) and generalized Borgen plot (gBp). Column 5: Hausdorff distances of classical Borgen plot (Bp) and the generalized Borgen plot (gBp). Column 6: Hausdorff distances of of the AFS computed by polygon inflation (PI) and the generalized Borgen plot (gBp).
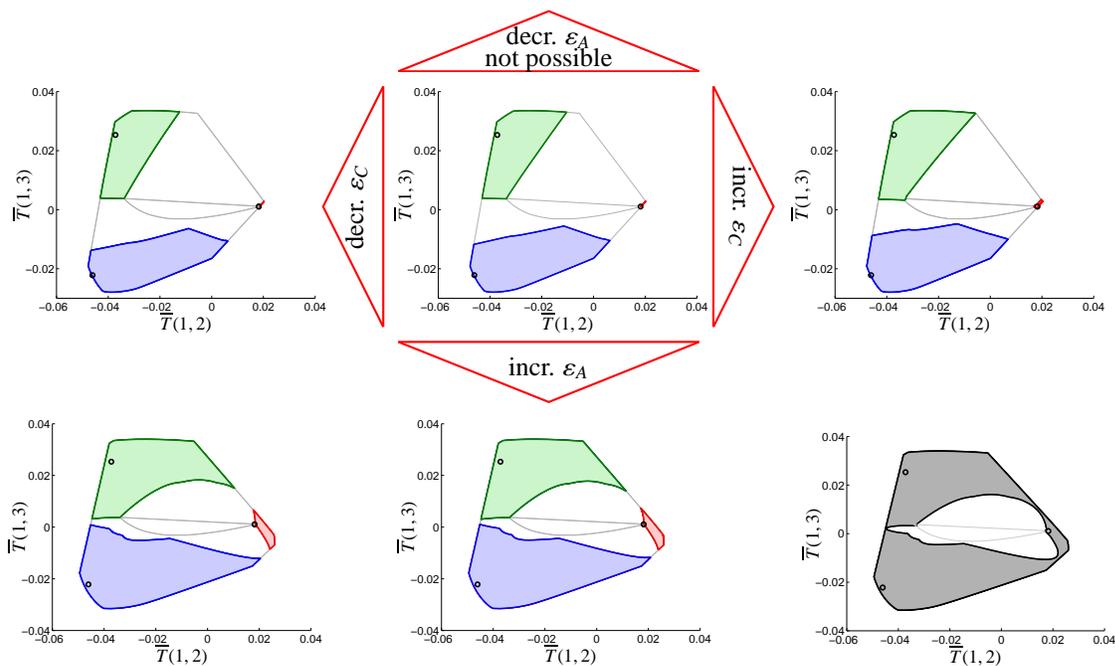
15

Figure 6: Parameter analysis for varying ε_C and ε_A for the FT-IR spectroscopic data from the hydroformylation process. The AFS is shown for for $\varepsilon_A = 0.000101$ in the first row and $\varepsilon_A = 0.00017$ in the second row. With the minimal entry $\min(D) = -0.000101$ the parameter ε_A must be equal or greater than $-\min(D) = 0.000101$; otherwise the geometric construction of the AFS is impossible. The parameter ε_C is set as follows. First column: $\varepsilon_C = 0$. Second column: $\varepsilon_C = 0.003$. Third column: $\varepsilon_C = 0.015$. The three black circles in each AFS indicate the positions of the true component spectra. These are the spectra of the olefin component, of the acyl complex and of the hydrido complex.

	$\varepsilon_C = 0$	$\varepsilon_C = 0.003$	$\varepsilon_C = 0.015$
$\varepsilon_A = 0.000101$	5.03s	5.44s	5.67s
$\varepsilon_A = 0.00015$	4.46s	5.39s	6.13s
$\varepsilon_A = 0.0002$	4.33s	5.52s	6.00s

Table 2: Computation times for the generalized Borgen plots applied to the hydroformylation spectral data from [11]. Further parameters are $\phi = 0.1°$ and $d = 0.001$, see part II of this paper.

December 2014) of the *FAC-PACK* software [22, 23]. The homepage of this software is

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

The *FAC-PACK* software allows to compute the AFS for two- and three-component systems. The core of this program code are implementations of the polygon inflation algorithm and the inverse polygon inflation algorithm [21]. The program includes also functionalities for the reduction of the AFS by the complementarity theorem, see [4, 20, 24]. The polygon inflation algorithm is an efficient tool for AFS computations for three-component systems. The new revision 1.2 of *FAC-PACK* includes the tangent- and line-moving algorithms for the construction of generalized Borgen plots for three component systems. The AFS can be computed with respect to the RS- or to the FSV-scaling. All results can easily be exported and can be compared with the purely numerical algorithm of polygon inflation.

5. Conclusion

So far the geometric-constructive Borgen plots for three-component systems were to some extent limited to model data and were not able to respond to the challenges of experimental spectroscopic data sets. Perturbed spectral data, low rank approximations and spectral data with small negative entries, e.g. from background subtraction, were not in the scope of Borgen plots. Generalized Borgen plots overcome these limitations. The mathematical fundamentals of such generalized Borgen plots have been presented - the pivotal point is the weakening of convex combinations towards affine combinations. New algorithms of computational geometry for the construction of generalized Borgen plots for three-component systems have been developed. These algorithms have already been published in the Borgen plot module of the *FAC-PACK* software. The detailed explanation of these geometric constructions is contained in a forthcoming second part of this paper. An

extension of these algorithms to four-component systems appears to be possible but can require very high computational times. Additional work on this is under progress.

We hope that the now enlarged range of applications of Borgen plots stimulates future developments of such global methods in chemometrics. These global methods make available the whole range of feasible solutions with a minimum of additional assumptions on the reaction system.

References

[1] H. Abdollahi, M. Maeder, and R. Tauler. Calculation and Meaning of Feasible Band Boundaries in Multivariate Curve Resolution of a Two-Component System. *Anal. Chem.*, 81(6):2115–2122, 2009.

[2] M. Akbari and H. Abdollahi. Investigation and visualization of resolution theorems in self modeling curve resolution (SMCR) methods. *J. Chemom.*, 27(10):278–286, 2013.

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

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

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

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

[7] 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 model-free analyses of multivariate data. Technical report, Technical report, 2014.

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

[9] 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.

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

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

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

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

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

[15] R. Rajkó. Some surprising properties of multivariate curve resolution-alternating least squares (MCR-ALS) algorithms. *J. Chemom.*, 23(4):172–178, 2009.

[16] R. Rajkó. Studies on the adaptability of different Borgen norms applied in self-modeling curve resolution (SMCR) method. *J. Chemom.*, 23(6):265–274, 2009.

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

[18] 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.

[19] 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.

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

[21] 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.

[22] 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.

[23] M. Sawall and K. Neymeyr. FAC-PACK: A software for the computation of multi-component factorizations and the area of feasible solutions, Revision 1.1. FAC-PACK homepage: http://www.math.uni-rostock.de/facpack/, 2014.

[24] 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.

[25] M. Sawall, N. Rahimdoust, C. Kubis, H. Schroder, D. Selent, D. Hess, H. Abdollahi, R. Franke, Börner A., and K. Neymeyr. Soft constraints for reducing the intrinsic rotational ambguity of the area of feasible solution. Technical report, University of Rostock, 2015. Technical Report.

[26] R. Tauler. Calculation of maximum and minimum band boundaries of feasible solutions for species profiles obtained by multivariate curve resolution. *J. Chemom.*, 15(8):627–646, 2001.