On hidden rank deficiency in MCR problems

Tomass Andersons[a], Mathias Sawall[a], Martina Beese[a,b], Christoph Kubis[b], Klaus Neymeyr[a,b]

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

Abstract

Pure component decomposition problems in chemometrics can be classified into rank-regular and rank-deficient problems. Rank-deficient problems are characterized by a spectral data matrix that has a lower rank than the number of chemical species. However, it is possible that there exists rank-regular factorization of the spectral data matrix, but none of these solutions can be interpreted chemically, and only a solution of the MCR problem with rank deficiency is chemically meaningful. Then we say that the underlying problem suffers from a hidden rank deficiency. In this paper, MCR problems with hidden rank deficiency are introduced and analyzed with several examples for problems of rank 2 and rank 3. The area of feasible solutions is determined with the help of additional constraints on the solution.

Keywords: hidden rank deficiency, multivariate curve resolution, pure component decomposition, nonnegative matrix factorization, area of feasible solutions, spectral recovery.

1. Introduction

Multivariate Curve Resolution (MCR) in chemometrics deals with low-dimensional pure component decompositions of a spectral data matrix $D \in \mathbb{R}_+^{k\times n}$ consisting of k measurements over n wave numbers. Most methods work with problems where the number of chemical components is equal to the rank of the matrix D, and aim to find one or all feasible nonnegative matrices C and S with

$$D = CS^T,$$

where C contains the pure component concentration profiles and S contains the spectra.

The ambiguity of the solutions can be represented with the help of the area of feasible solutions (AFS) [5, 17]. Some standard methods for finding the AFS are based on geometric construction [15, 5, 20] or on the solution of optimization problems [24, 11, 17].

Rank-deficient problems require a different approach than rank-regular problems because the concentration factor C (usually) has a lower rank than the spectral factor S due to linear dependency of the concentration profiles. For rank-deficient problems, the AFS of the matrix C can be found by numerical methods, e.g., polygon inflation [25], or geometric methods [23]. Furthermore, the band boundaries of the spectral factor can also be approximated for known concentration profiles [22]. In experimental setups, rank deficiency can sometimes be resolved by rank augmentation, see, e.g., [2, 21].

But in between these types of problems there exists another distinct situation, where one or more components remain hidden to standard analysis methods. This can happen for any data set. In this case, the data set appears to be rank-regular, since there exists a nonnegative matrix factorization of D with the correct rank. However upon closer inspection, it is evident that no chemically meaningful factorization is possible with this rank. The rank deficiency cannot be detected only from the rank and the nonnegative rank of D (definitions of these rank notions are given in Sec. 2) and is therefore hidden. Finally, the concept of constrained rank is introduced. With its help, the AFS can be approximated. In the following section, we analyze a simple model data set to illustrate a possible misinterpretation of a problem with a hidden rank deficiency.

1.1. Problem misconception

A short example with a given reaction kinetics and abstract pure component spectra is constructed to demonstrate how a rank-deficient problem can be misinterpreted. The methods that are used to analyze the problem are explained in detail in the following sections of this work. The example is as follows.

Example 1.1. *The factor $C \in \mathbb{R}^{101\times 4}$ is computed for the kinetic*

$$\boldsymbol{A} \xrightarrow{k_1} \boldsymbol{B}, \qquad \boldsymbol{C} \xrightarrow{k_2} \boldsymbol{D} \tag{1}$$

with the kinetic constants $k_1 = 1$ and $k_2 = 2$ and the initial concentrations $C_{1,:} = (1, 0, 1, 0)$ on an equidistant grid with $k = 101$ nodes for the time interval $t \in [0, 5]$. The factor S is defined with shifted Gaussians as

$$s_i(x) = a_i \exp\left(-\frac{(x - c_i)^2}{d_i}\right) + b_i, \qquad \text{for } i = 1, \ldots, 4,$$

with $c_1 = 20$, $c_2 = 40$, $c_3 = 60$, $c_4 = 80$, $b_1 = 0.02$, $b_2 = 0.01$, $b_3 = 0.005$, $b_4 = 0.001$ and $a_i = 1$, $d_i = 100$ for $i = 1, \ldots, 4$ (where $(1, 2, 3, 4) = (\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{D})$) on an equidistant grid with $n = 100$ nodes for $x \in [0, 99]$. The pure component profiles are shown in Fig. 1. The rank and the nonnegative rank of the matrix D is 3, but the number of chemical components is 4.

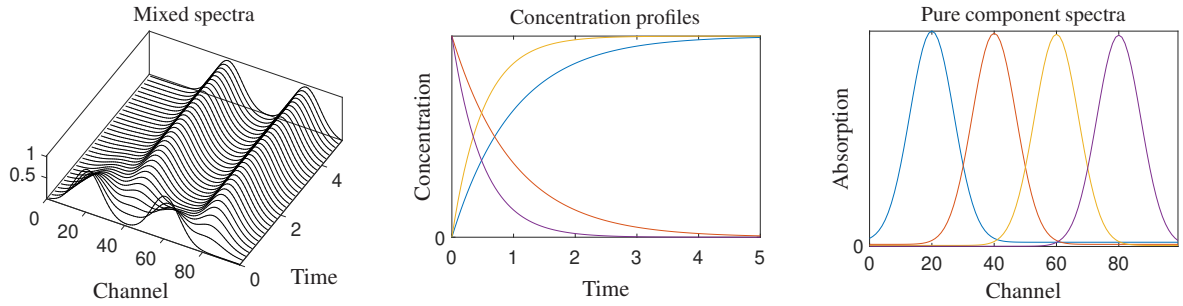

Figure 1: The data matrix and pure component factors for the model problem from Example 1.1. It holds that $\operatorname{rank}(D) = \operatorname{rank}_+(D) = 3$, but the problem is rank-deficient because four chemical components are expected.

Next, the AFS is computed for factor C, see, e.g., [5, 20, 24] for possible techniques. This set is not empty because there exist nonnegative factors with three components, which is usually expected for problems of rank 3. The associated feasible concentration profiles are computed from the AFS and presented in Fig. 2 together with the true profiles. As can be seen in Fig. 3, only two of the true profiles are included in the AFS. Furthermore, the results lead to a complete misconception of the underlying reaction kinetics. The solution suggests a three-component system

$$\mathbf{X} \to \mathbf{Y} \to \mathbf{Z}. \tag{2}$$

Such a solution with an intermediate species $\mathbf{Y}$ is easily representable within the space of true concentration profiles of the reaction scheme (1). The concentration profiles $c_{\mathbf{A}}(t) = c_0 \exp(-k_1 t)$ and $c_{\mathbf{C}}(t) = c_0 \exp(-k_2 t)$ with the initial concentration values $c_0 = 1$ of (1) allow the linear combination $c(t) = \exp(-k_1 t) - \exp(-k_2 t)$. This profile $c(t)$ has the typical shape of a potential intermediate species $\mathbf{Y}$ with $c(0) = 0$, $\lim_{t\to\infty} c(t) = 0$ and its positive maximum at $t_{\max} = \ln(k_2/k_1)/((k_2 - k_1)$, which for the given values of $k_2 = 2$ and $k_1 = 1$ equals $t_{\max} = \ln(2) \approx 0.69$. Such a result is approximately consistent with Fig. 2. All this suggests that the consecutive reaction scheme (2) with three chemical species is the true scheme behind the given data set. Analysis with MCR-ALS (see, e.g., [27]) leads to similar conclusions. It is a remarkable fact that it is not possible to identify the true kinetics without additional knowledge on the reaction system.

To account for the number of components in Example 1.1, an additional constraint on the solution is needed. Here, the monotonicity of all concentration profiles is used to restrict the AFS. The AFS and the generalized area of feasible solutions with respect to hidden rank deficiency (denoted Hidden-AFS) are plotted in Fig. 3. It can be observed that a nested triangle can be constructed between the inner and the outer polygons. This is a geometric indication that a three-component factorization exists, see [5, 10]. However, by considering the monotonicity constraint on the outer polygon, no triangle can be constructed in the feasible regions, i.e., no solution with three components and monotone concentration profiles is possible. Hidden rank deficiency can also be indicated by the existence of a known profile that is not in the computed AFS for three components.

The feasible concentration profiles resulting from the Hidden-AFS with the monotonicity constraint are shown in Fig. 4 and it can be observed, that the true solution is included.

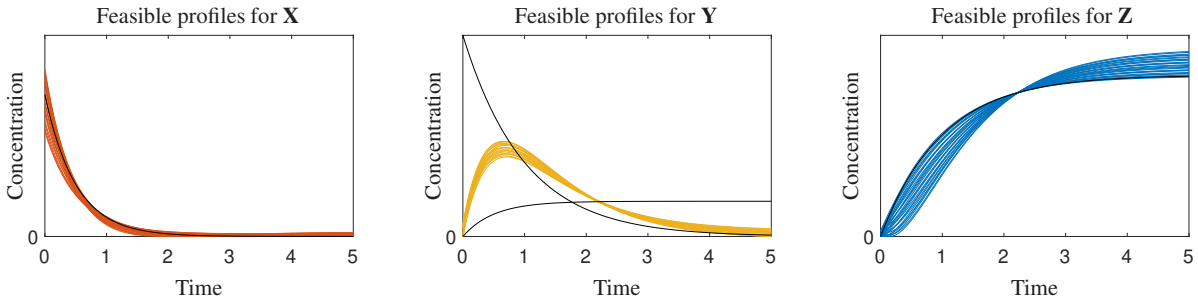

Figure 2: The feasible bands for the model problem from Example 1.1 without any additional knowledge on the solution. The true profiles are shown as black lines. Since $\mathrm{rank}(D) = \mathrm{rank}_+(D) = 3$ the AFS is not empty. But the feasible profiles for **Y** are not chemically meaningful for the kinetics. This demonstrates how rank deficiency can lead to a wrong solution. The feasible concentration profiles with respect to the monotonicity constraint and rank deficiency are presented in Fig. 4.

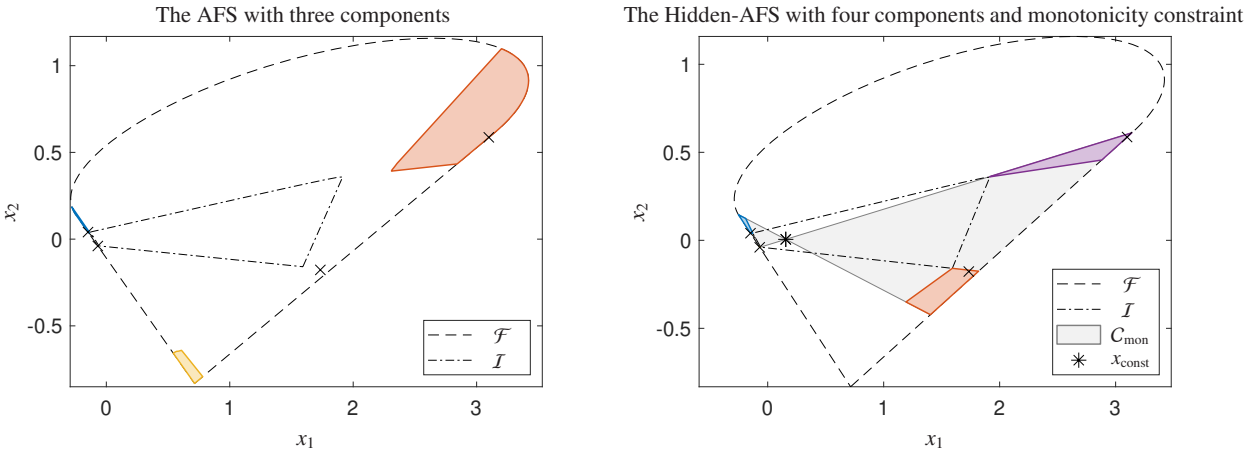

Figure 3: The AFS and the Hidden-AFS for the factor C for the Example 1.1. Left: The AFS leads to wrong conclusions. Not all of the points representing the true solution (×) are included in the AFS. Right: The Hidden-AFS with respect to the monotonicity constraint. Each of the four isolated subsets represents the ambiguity of a concentration profile of a pure component. The subset with the approximate coordinates $(-0.07, -0.04)$ is very small. The area of monotone profiles C_{mon} also includes the subsets of the Hidden-AFS and x_{const} shows the low-dimensional representation of a constant profile.

1.2. Overview

This paper is organized as follows. In Sec. 2, we review some important concepts on rank deficiency in MCR problems and introduce the notion of constrained rank to define the hidden rank deficiency. Furthermore, the Hidden-AFS is analyzed without and with the constraints of monotonicity and unimodality. Sec. 3 illustrates important concepts with the help of rank-2 and rank-3 model problems and discusses different approaches to solve them. An experimental data set is analyzed in Sec. 4.

1.3. List of symbols

$\mathbb{R}_+$	the set of nonnegative real numbers,
rank_+	nonnegative rank,
rank_f	constrained rank,
D	spectral data matrix,
C	matrix with pure component concentration profiles in columns,
S	matrix with pure component spectral profiles in columns,
k	number of spectra,

n	number of spectral channels,
s	rank of matrix D,
m	nonnegative rank of matrix D,
r	number of chemical components,
$U\Sigma V^T$	the truncated singular value decomposition of D,
$\mathcal{F}$	the outer polygon or polytope (FIRPOL),
$\mathcal{I}$	the inner polygon or polytope (INNPOL),
$\mathcal{M}$	AFS for rank-regular problems,
$\mathcal{N}$	AFS with respect to (open) rank deficiency,
$\mathcal{N}_\mathrm{h}$	Hidden-AFS with respect to hidden rank deficiency,
$\mathcal{N}_\mathrm{constr}$	Hidden-AFS with respect to hidden rank deficiency and additional constraints,
C_mon	area of monotone profiles,
$s^\mathrm{sc}_\mathrm{const}$	a scaled constant profile,
x_const	the coordinates of the low-dimensional representation of $s^\mathrm{sc}_\mathrm{const}$.

2. Theoretical background to hidden rank deficiency

2.1. The rank-deficient MCR problem

This paper considers a special case of the MCR problem

$$D = CS^T + E,$$

where $D \in \mathbb{R}^{k\times n}$ with $\mathrm{rank}(D) = s$ is given. For this problem, all feasible nonnegative factors $C \in \mathbb{R}^{k\times r}_+$ and $S \in \mathbb{R}^{n\times r}_+$ are sought after, while the error matrix E is as small as possible. The variable r is also understood as the number of pure components underlying the data matrix. Rank deficiency means that $r > s$. Measurement errors in $E \in \mathbb{R}^{k\times n}$ are disregarded for the theoretical part of this paper, but are treated in Sec. 4.

The term rank is sometimes used ambiguously in chemometrics. In linear algebra the rank of a matrix is defined as the maximal number of linearly independent columns in the matrix. However, in floating-point arithmetic this definition would almost always result in full rank. Therefore, in these cases, the term "rank" is implicitly understood as *numerical rank*, i.e., only those singular values of D are accepted that are greater than a certain tolerance value, see [12]. Furthermore, experimental data matrices are usually affected by noise, and in chemometrics "rank" often means the number of singular values that do not correspond to noise. This is also known as the *chemical rank*. The determination of the chemical rank is a well-known problem, see, e.g., [16]. In this paper we also follow the usual convention, i.e., we use numerical rank for simulated noise-free data and chemical rank for experimental data. These concepts are also discussed in [18].

Rank-deficient problems are usually characterized by their *nonnegative rank*, denoted by rank_+. The nonnegative rank is equal to the smallest possible dimension of the nonnegative factors C and S. Hence, if we can find nonnegative matrices $C \in \mathbb{R}^{k\times m}_+$ and $S \in \mathbb{R}^{n\times m}_+$ so that $D = CS^T$ holds, but there are no such matrices $C \in \mathbb{R}^{k\times m-1}_+$ and $S \in \mathbb{R}^{n\times m-1}_+$, then $\mathrm{rank}_+(D) = m$, see [6, 10, 9].

Note, that throughout this paper we also assume that either $\mathrm{rank}(C) = \mathrm{rank}(D)$ or $\mathrm{rank}(S) = \mathrm{rank}(D)$. In mathematics this additional condition to the nonnegative rank is known as the restricted nonnegative rank, see, e.g., [10, 9].

Furthermore, $s = \mathrm{rank}(D) \leq \mathrm{rank}_+(D) = m$ follows from the nonnegativity condition in the definition. This inequality and the dimension r of the factors C and S helps us to categorize MCR problems.

1. **Rank-regular** MCR problems are described by the equation

   $$s = \mathrm{rank}(D) = \mathrm{rank}_+(D) = m = r,$$

   where the rank of the data matrix is equal to the nonnegative rank and is equal to the number of chemical components.

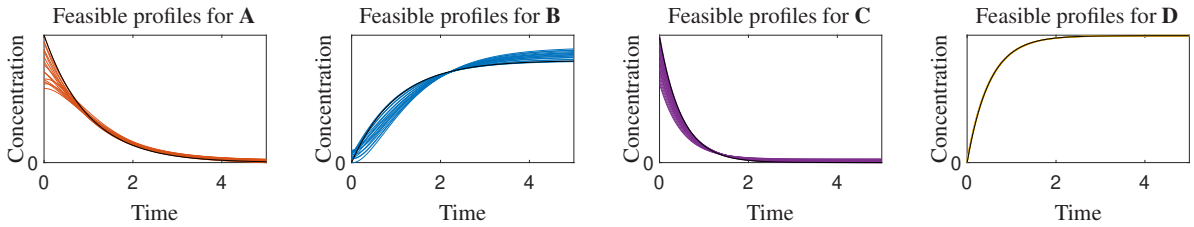

Figure 4: The feasible bands for the model problem from Example 1.1 with respect to the monotonicity constraint and rank deficiency. Unlike in Fig. 2, the true profiles (black lines) are included in the feasible bands.

2. **(Open) rank-deficient** MCR problems are represented by the strict inequality in

$$s = \operatorname{rank}(D) < \operatorname{rank}_+(D) = m = r.$$

In this case, it is not possible to find two nonnegative matrices C and S with s columns so that $D = CS^T$.

However, it seems that this classification does not completely describe all possible types of MCR problems. This definition disregards whether the found solution is chemically meaningful. In fact, the dimension r of the factors C and S depends neither on the rank nor on the nonnegative rank of D, but on the number of meaningful chemical components in the bilinear spectroscopic data set. Therefore, it is necessary to define a measure to describe this property.

Definition 2.1 (Constrained rank)**.** *The constrained rank of $D \in \mathbb{R}^{k\times n}$, denoted by* $\operatorname{rank}_f(D)$*, is the minimum integer r so that there exist factors $C \in \mathbb{R}^{k\times r}$ and $S \in \mathbb{R}^{n\times r}$ with $D = CS^T$ and these factors satisfy some constraint function $f(C, S) = 0$.*

The nonnegativity of the factors C and S is the most important constraint in the context of MCR problems and in this paper we assume that the nonnegativity is always one of the constraints in the function $f(C, S)$. This means, that nonnegative rank is a special case of the constrained rank. Other constraints can be applied simultaneously, for example, monotonicity for all profiles of one factor. Then it follows that $m = \operatorname{rank}_+(D) \leq \operatorname{rank}_f(D)$ and in an optimal case $\operatorname{rank}_f(D) = r$. These relationships lead to a categorization of the underlying problems.

3. **Hidden rank-deficient** MCR problems are represented by the strict inequality in

$$s = \operatorname{rank}(D) \leq \operatorname{rank}_+(D) = m < r.$$

This means, that the nonnegative factorization with the least number of columns m is not a chemically meaningful one. If $\operatorname{rank}_f(D) = r$, then the constraint function f helps to resolve the problem with the correct number of chemical components.

This is in contrast to MCR problems with open rank deficiency or rank regular MCR problems, where the equality $m = \operatorname{rank}_+(D) = \operatorname{rank}_f(D) = r$ means that the factors with the smallest number of columns have a chemically meaningful interpretation. For MCR problems with hidden rank deficiency, the constrained rank is equal to the number of chemical components if the constraint function $f(C, S)$ is chosen correctly.

Direct computation of the constrained rank from the given spectral matrix D alone is not feasible and depends heavily on the chosen constraints. This is particularly difficult for experimental data sets with perturbations, as discussed in Sec. 4. The constraint function $f(C, S)$ should be applied in the context of numerical or chemical rank and should have some tolerance.

2.2. AFS as a representation of solutions

The solution of an MCR problem is usually not unique. This is known as rotational ambiguity, see, e.g., [1]. Therefore, it is of interest to find the AFS, which includes the variety of factorizations up to scaling and permutations.

One possible approach uses the singular value decomposition (SVD) of the spectral data matrix D. In particular, we denote the truncated SVD of D with $U\Sigma V^T$, where $U \in \mathbb{R}^{k\times s}$, $\Sigma \in \mathbb{R}^{s\times s}$ and $V \in \mathbb{R}^{n\times s}$.

Here, the main concepts for the construction of the AFS are repeated for both rank-regular problems (see, e.g., [17]) and open rank-deficient problems (see [25]), since hidden rank deficiency is related to both types of problems. The main geometric objects that reflect the nonnegativity constraints of the pure components are the outer polytope $\mathcal{F}$ (also known as FIRPOL)

$$\mathcal{F} = \left\{x \in \mathbb{R}^{s-1} : (1, x^T)V^T \geq 0\right\}$$

and the inner polytope $\mathcal{I}$ (also known as INNPOL)

$$\mathcal{I} = \text{convhull}(\{a_i \in \mathbb{R}^{s-1} : i = 1, \ldots, k\})$$

which is spanned by the data representing points (or data points)

$$a_i = \left(\frac{(U\Sigma)_{i,2:s}}{(U\Sigma)_{i,1}}\right)^T.$$

Here and further, the colon notation is used for vectors and matrices, see also [12] or MatLab notation. Furthermore, for the ease of representation it is assumed that the rank deficiency is caused by the factor S, i.e., $\text{rank}(S) = \text{rank}(D)$ and this is used for the theory in this section. Rank deficiency in MCR problems is most likely induced by the factor C due to a linear dependency of the concentration profiles. However, sometimes the analysis is more convenient by assuming that S is the rank-deficient factor, see [25].

Remark 2.2. *A transposition of $D = CS^T$ results in $D^T = SC^T$ where the factors C and S have changed their positions and are transposed. Therefore, it does not make a significant difference, whether C or S causes rank deficiency. We can flip between both cases by applying the analysis to D or D^T. This gives the justification to assume that the factor S is responsible for the rank deficiency in the following analysis. The model data sets in Sec. 3 and the experimental data set in Sec. 4 are analyzed by using D^T.*

The factor S is reconstructed from the SVD with the help of a transformation matrix $T \in \mathbb{R}^{r\times s}$ (for rank regular problems $r = s$, for open rank-deficient problems $r = m$) as $S = TV^T$. This can be used to find a low-dimensional representation of all feasible solutions of the MCR problem, since it is contained in the rows of T. For this purpose the matrix T is written as

$$T = \begin{pmatrix} 1 & x^T \\ \mathbf{1} & W \end{pmatrix}$$

where $\mathbf{1} = (1, \ldots, 1)^T$ and $W \in \mathbb{R}^{(r-1)\times(s-1)}$ under the weak assumption of irreducibility of D^TD (see [17]). Each row of the matrix T represents a profile in the $(s-1)$-dimensional space. Geometrically, the matrix T is feasible if and only if the polytope that is spanned from the rows of $T_{:,2:s}$ is nested between $\mathcal{I}$ and $\mathcal{F}$, see, e.g., [5, 20, 9]. A proof of this is also given in Thm. 4.6 in [17] for rank-regular problems and in Thm. 4.2 in [25] for MCR problems with (open) rank deficiency.

This motivates the definition of the AFS. Each point in the AFS represents one spectral profile of a pure component, for which a full factorization is possible. The AFS for rank-regular problems is

$$\mathcal{M} = \left\{x \in \mathbb{R}^{s-1} : \text{ exists } T \in \mathbb{R}^{s\times s} \text{ with } T_{1,2:s} = x^T,\ T_{:,1} = \mathbf{1},\ \text{rank}(T) = s,\ U\Sigma T^{-1} = C \geq 0,\ TV^T = S^T \geq 0\right\}$$

and for open rank-deficient problems

$$\mathcal{N} = \left\{x \in \mathbb{R}^{s-1} : \text{ exists } T \in \mathbb{R}^{m\times s},\ C \in \mathbb{R}_+^{k\times m} \text{ with } T_{1,2:s} = x^T,\ T_{:,1} = \mathbf{1},\ TV^T = S^T \geq 0,\ CS^T = D\right\}.$$

2.3. Hidden-AFS for problems with hidden rank deficiency

The low-dimensional representation of the AFS is a useful analytical tool for rank-regular and open rank-deficient problems. A similar approach is of interest for the case of hidden rank deficiency, i.e., a definition of Hidden-AFS. It is possible to pursue a similar definition as before and use the number of components r, i.e.,

$$\mathcal{N}_\text{h} = \left\{x \in \mathbb{R}^{s-1} : \text{ exists } T \in \mathbb{R}^{r\times s},\ C \in \mathbb{R}_+^{k\times r} \text{ with } T_{1,2:s} = x^T,\ T_{:,1} = \mathbf{1},\ TV^T = S^T \geq 0,\ CS^T = D\right\}.$$

As before, $r > m$ implies the case of hidden rank deficiency.

However, in this case some vertices in the low-dimensional representation can be located freely inside $\mathcal{F}$, because there exists a polytope with fewer vertices, compare this to the nonnegative rank in [10]. A vertex of the polytope can even be located inside of $\mathcal{I}$. This is proved by the following theorem.

Theorem 2.3 (Hidden-AFS for problems with hidden rank deficiency)**.** *The hidden-AFS of the MCR problem with hidden rank deficiency equals the outer polytope*

$$\mathcal{N}_{\mathrm{h}} = \mathcal{F}.$$

Proof. If a problem has a hidden rank deficiency, then we can assume that there exists a rank regular solution or a solution with open rank deficiency, i.e., there exists a $T \in \mathbb{R}^{m\times s}$ and $C \in \mathbb{R}_+^{k\times m}$ with $T_{:,1} = \mathbf{1}$, $TV^T = S^T \geq 0$ and $CS^T = D$.

The goal is to construct such a $\widehat{T} \in \mathbb{R}^{r\times s}$ that there exists $\widehat{C} \in \mathbb{R}_+^{k\times r}$ with $\widehat{T}_{:,1} = \mathbf{1}$, $\widehat{T}V^T = \widehat{S}^T \geq 0$ and $\widehat{C}\widehat{S}^T = D$. Due to hidden rank deficiency $r > m$. We assume w.l.o.g. that $r = m + 1$. If r is larger, then any further rows can be constructed analogously.

In order to show that any point $x \in \mathcal{F}$ of the $m + 1$ row of $\widehat{T}$ is contained in $\mathcal{N}_{\mathrm{h}}$, we set

$$\widehat{C} = (C, \mathbf{0}) \quad \text{and} \quad \widehat{S} = \left(S, V\begin{pmatrix}1\\ x\end{pmatrix}\right).$$

Then $\widehat{S} \geq 0$ because $S \geq 0$ and $V(1, x)^T \geq 0$. The latter follows from $x \in \mathcal{F}$. Furthermore,

$$\widehat{C}\widehat{S}^T = CS^T + \mathbf{0}\left(V\begin{pmatrix}1\\ x\end{pmatrix}\right)^T = CS^T = D$$

and this completes the proof. □

Remark 2.4. *The proof of Thm. 2.3 uses a trivial choice of $\widehat{C}_{:,m+1} = \mathbf{0}$. This is usually not a satisfactory solution. However, there can (and usually do) exist non-trivial solutions. To achieve this $\widehat{S}_{:,m+1}$ is written as linear combination*

$$\widehat{S}_{:,m+1} = \sum_{i=1}^{m} \beta_i S_{:,i}, \quad \beta_i \in \mathbb{R}$$

because $\operatorname{rank}(S) = \operatorname{rank}(\widehat{S}) = s$*. The only constraint on the values of β_i is the nonnegativity of $\widehat{S}_{:,m+1}$. Therefore, $\widehat{T}_{m+1,:}$ is constrained only by $\mathcal{F}$. It remains to be shown that $\widehat{S}$ leads to a feasible nonnegative factorization which is given as*

$$\begin{aligned}
\widehat{C}\widehat{S}^T &= \sum_{i=1}^{m+1} \widehat{C}_{:,i}(\widehat{S}_{:,i})^T \\
&= \left(\sum_{i=1}^{m} \widehat{C}_{:,i}(S_{:,i})^T\right) + \widehat{C}_{:,m+1}\sum_{i=1}^{m} \beta_i (S_{:,i})^T \\
&= \sum_{i=1}^{m} \left(\widehat{C}_{:,i} + \beta_i \widehat{C}_{:,m+1}\right)(S_{:,i})^T \\
&= \sum_{i=1}^{m} C_{:,i}(S_{:,i})^T \\
&= D.
\end{aligned}$$

This implies that if $C_{:,i} = \widehat{C}_{:,i} + \beta_i \widehat{C}_{:,m+1}$ for $i = 1, \ldots, m$, then $\widehat{C}\widehat{S}^T = D$. Finally, we set

$$\widehat{C}_{:,i} = C_{:,i} - \beta_i \widehat{C}_{:,m+1}$$

and the feasible entries of $\widehat{C}_{:,m+1}$ *are given by an interval*

$$\widehat{C}_{h,m+1} \in [0, \min_{i=1,\dots,m}(\gamma_i)], \quad \text{with}$$
$$\gamma_i = \begin{cases} 0 & \beta_i < 0 \\ \frac{1}{\beta_i} C_{h,i} & \beta_i > 0 \end{cases} \quad \text{and}$$
$$\gamma_i > 0 \quad \text{if } \beta_i = 0,$$

for each $h = 1, \dots, k$ *because this choice also leads to* $\widehat{C} \geq 0$.

Thm. 2.3 implies that the definition of $\mathcal{N}_\mathrm{h}$ is practically meaningless, since the same information is already included in the outer polytope. That is, for problems with hidden rank deficiency $\mathcal{F}$ stands for all nonnegative profiles for the factor S that can be constructed from the SVD.

These results are unsatisfactory because the ambiguity represented by the Hidden-AFS is very large in this case. However, some solutions can be disregarded if additional information is provided. In general, problems with hidden rank deficiency require additional constraints for a successful analysis. These constraints should reflect those imposed by the constrained rank, and then a different variant of the Hidden-AFS can be formally defined.

Definition 2.5. *Let* $f(C,S)$ *be a constraint function of the factors* C *and* S*, where* $f(C,S) = 0$ *if the constraints are satisfied. Further, let* $\mathrm{rank}(D) = s$ *and* $\mathrm{rank}_f(D) = r$*. The Hidden-AFS of an MCR problem with hidden rank deficiency and additional constraints is defined as*

$$\mathcal{N}_\mathrm{constr} = \left\{ x \in \mathbb{R}^{s-1} : \text{ exists } T \in \mathbb{R}^{r\times s},\ C \in \mathbb{R}^{k\times r} \text{ with } T_{1,2:s} = x^T,\ T_{:,1} = \mathbf{1},\ TV^T = S^T,\ CS^T = D,\ f(C,S) = 0 \right\}.$$

The chosen constraint function $f(C,S)$ must be strong enough to prevent a factorization with less pure components, i.e., the condition $\mathrm{rank}_f(D) = r$ is important. Consequently, there should exist no polytope with less than r vertices. Then the methods from Sec. 3.2 and 3.3 can be applied. However, this is not always possible, see Sec. 2.4, 3.5 and 3.6.

The implementation of the constrained rank in the computation of a factorization or the Hidden-AFS is not easy. In contrast to the precisely defined AFS $\mathcal{M}$ and the AFS for rank-deficient problems $\mathcal{N}$, the Hidden-AFS for a problem with hidden rank deficiency, is much more abstract, since there is no mathematical definition of a chemically meaningful factorization. The results strongly depend on the chosen constraints for the profiles of the MCR problem. This paper focuses on a classic constraint in chemometrics, namely, monotonicity. However, other constraints can also be used, see [7, 8, 4, 3] for some examples.

2.4. Properties of the Hidden-AFS

The proof of Thm. 2.3 leads to the following remark on the polytope that represents a feasible solution.

Remark 2.6. *The low-dimensional representation of a feasible solution is given by points that must be located in* $\mathcal{F}$*. A subset of these points must form a polytope that encloses* $\mathcal{I}$*. The remaining points can be located anywhere in* $\mathcal{F}$*.*

In the proof of Thm. 2.3 a solution with less than r components is used. This is useful if only the nonnegativity constraint can be applied. Then $\mathcal{N}_\mathrm{h}$ describes the problem instead of $\mathcal{N}_\mathrm{constr}$, because $\mathrm{rank}_f(D) = r$ is not satisfied. Therefore, sometimes even $\mathcal{M}$ or $\mathcal{N}$ can be considered for a part of the solution, see Sec. 3.5 or 3.6. However, $\mathcal{M}$ or $\mathcal{N}$ generally do not describe all solutions in this case and can only be applied with additional knowledge about the underlying problem.

Furthermore, the Hidden-AFS is bounded. It follows from Thm. 2.3 that the Hidden-AFS is a subset of $\mathcal{F}$ and $\mathcal{F}$ itself is bounded, see for example Thm. 4.2 in [17]. This is an analogous result to the one in the case of open rank deficiency, see Remark 4.1 in [25].

Finally, the constraint of $\mathcal{N}_\mathrm{constr}$ applies only to vertices and not the whole polytope that represents a solution.

Remark 2.7. *Only the vertices of the polytope must be located in the regions, where the constraint is satisfied. The edges of the polytope must satisfy the condition for* $\mathcal{I}$*, i.e., they must enclose it, but they are not associated with the constraint function.*

2.5. Monotonicity constraint

Monotonicity is a particularly useful constraint in reducing the ambiguity of feasible solutions, see also [29, 26]. Monotone behavior can be observed for the concentration profiles of reactants and products. According to Remark 2.2 we impose the nonnegativity constraint on the factor S and it leads to the outer boundary $\mathcal{F}$. By allowing only monotone increasing or monotone decreasing profiles in the rank-deficient factor S, the feasible regions of the low-dimensional representation are further reduced, see Fig. 5.

Regions that represent profiles satisfying the monotonicity constraint are numerically determined by a pointwise evaluation of the constraint function. The outer polytope $\mathcal{F}$ is covered by a fine grid and the corresponding profile of each grid point is tested. The monotonicity constraint is applied as described in [29, 26]. In summary, a squared sum of all deviations from monotonicity is calculated. Furthermore, if several consecutive points of a profile have an incorrect trend, each of them is compared to the last point with the correct monotonicity behavior. For experimental data sets, small deviations from monotonicity are tolerated. For the following statements, we assume that $s = \operatorname{rank}(D) = \operatorname{rank}(S)$ (see Remark 2.2). We also assume that the data set comes from a closed chemical system with one or more mass balance constraints. The mass balance condition is also known as the closure constraint, see [19].

In general, numerical approaches for the computation of Hidden-AFS are used, see Sec. 3.2. However, for some noise-free cases a geometric calculation of the Hidden-AFS is possible, described in Sec. 3.3. Some theoretical considerations on monotonicity are required for this approach. The goal is to show that the monotonicity constraint is given by lines, which can be used for the geometric construction of the Hidden-AFS. First, we show a fundamental property for such chemical systems.

Lemma 2.8. *If the factor S contains concentration profiles (according to Remark 2.2) of a chemical system with a mass balance constraint, then in the corresponding low-dimensional representation there exists exactly one point $x_{\text{const}} \in \mathbb{R}^{s-1}$ representing a constant profile.*

Proof. First, the existence of x_{const} is shown. According to mass balance constraint, there exist coefficients a_i, $i = 1 \dots p$ so that the p original profiles of S result in a constant profile

$$a_1 S_{:,1} + \cdots + a_p S_{:,p} = s_{\text{const}},$$

where $s_{\text{const}} \in \mathbb{R}^n$ denotes a vector in with constant entries.

The low-dimensional representation requires that the original profiles are scaled so that the first singular vector has a coefficient 1, see Sec. 2.2. Thus,

$$\frac{a_1 b_1}{d} \underbrace{\left(\frac{S_{:,1}}{b_1}\right)}_{S^{\text{sc}}_{:,1}} + \cdots + \frac{a_p b_p}{d} \underbrace{\left(\frac{S_{:,p}}{b_p}\right)}_{S^{\text{sc}}_{:,p}} = \underbrace{\frac{s_{\text{const}}}{d}}_{s^{\text{sc}}_{\text{const}}} = \text{const},$$

where $S^{\text{sc}}_{:,i}$ represent a scaled profile by the scaling factor b_i. However, the constant vector s_{const} must also be scaled to have a low-dimensional representation. Is we use $d = a_1 b_1 + \cdots + a_p b_p$ as the scaling factor, then $s^{\text{sc}}_{\text{const}}$ is scaled so that the first singular vector contributes to the expansion with the coefficient 1.

The low-dimensional representation x_{const} of $s^{\text{sc}}_{\text{const}}$ is given by the low-dimensional representations x_i of the scaled profiles $S^{\text{sc}}_{:,i}$, i.e.,

$$\frac{a_1 b_1}{d}(1, x_1^T) + \cdots + \frac{a_r b_r}{d}(1, x_p^T) = (1, x_{\text{const}}^T).$$

The uniqueness of x_{const} is proven by contradiction. We assume, that there exists a different point $\hat{x}_{\text{const}}$ corresponding to the constant profile $s^{\text{sc}}_{\text{const}}$ which is unique due to scaling. Then

$$V_{:,1} + V_{:,2:p} x_{\text{const}}^T = s^{\text{sc}}_{\text{const}} = V_{:,1} + V_{:,2:p} \hat{x}_{\text{const}}^T$$

and

$$((x_{\text{const}})_1 - (\hat{x}_{\text{const}})_1) V_{:,2} + \cdots + ((x_{\text{const}})_{p-1} - (\hat{x}_{\text{const}})_{p-1}) V_{:,p} = 0.$$

From the linear independence of the singular vectors follows $x_{\text{const}} = \hat{x}_{\text{const}}$. □

Remark 2.9. *The low-dimensional representation of a constant profile is located at the origin (i.e., $x_{\text{const}} = \mathbf{0}$) if and only if the first right singular vector of D is a constant vector. This follows from $s^{\text{sc}}_{\text{const}} = V_{:,1} + V_{:,2:s} x^T_{\text{const}}$.*

Next, we prove a simple property for convex combinations of profiles.

Lemma 2.10. *The low-dimensional representation of a convex combination of two profiles is located on the line segment between the two respective low-dimensional representations.*

Proof. The convex combination of two scaled profiles results in a scaled profile $\lambda S^{sc}_{:,i} + (1-\lambda) S^{sc}_{:,j}$ with $\lambda \in [0,1]$, because the first singular vector of s^{sc} has the coefficient $\lambda + (1-\lambda) = 1$.

With the help of singular values this equality can also be written in the corresponding low-dimensional representation $\lambda x_i + (1-\lambda) x_j$ with $\lambda \in [0,1]$. This implies, that the line segment connecting the low-dimensional representations of the two profiles includes their convex combinations. □

Remark 2.11. *If the mass balance constraint holds then x_{const} is located in the convex hull of the low-dimensional representation of the true profiles. This follows from an extension of Lemma 2.10 to polytopes.*

A similar argumentation to Lemma 2.10 can be used to reduce the ambiguity if the mass balance constraint of some components is known, see Sec. 3.4 for an example. Finally, the following lemma proves the geometric structure of monotone subsets, see also Fig. 5.

Lemma 2.12. *If a point exists that corresponds to a monotone profile, then a connected subset with non-increasing concentration profiles and a connected subset with non-decreasing concentration profiles exist in the low-dimensional representation. The boundaries of both subsets are given by lines that intersect at the point of constant profiles.*

Proof. First, we show that the subsets are connected. The convex combination of two scaled profiles with the same monotonicity yields a scaled monotone profile. It follows from Lemma 2.10 that any profile that corresponds to a point on the line segment that connects the two scaled profiles in the low-dimensional representation is also located in the same subset. Therefore, the subsets with the same monotonicity must be connected.

Then we show that the boundaries of a monotonicity subset are given by lines. For this we consider the linear combination of $s^{\text{sc}}_{\text{const}}$ and a monotone profile $S^{\text{sc}}_{:,l}$ with a low-dimensional representation x_l. The resulting profile $s^{\text{sc}}_{\text{const}} + \lambda S^{\text{sc}}_{:,l}$ has to be scaled with the scaling factor $d = 1 + \lambda$

$$\frac{1}{d} s^{sc}_c + \frac{\lambda}{d} S^{sc}_{:,l},\ \lambda \in \mathbb{R}$$

so that its low-dimensional representation

$$\frac{1}{d} x_{\text{const}} + \frac{\lambda}{d} x_l,\ \lambda \in \mathbb{R}$$

lies on a line going through x_{const} and x_l. The monotonicity is reversed for negative values of λ. Therefore, the low-dimensional representations of all monotone profiles are located on lines that intersect in the point of constant profiles x_{const}. This proves the statement. □

The previous statements apply only to problems with negligible perturbations. Experimental data may show a different behavior, especially if deviations from monotonicity or mass balance are allowed. See, for example, Fig. 16.

2.6. Unimodality constraint

Unimodality is another effective constraint to reduce the ambiguity of the solutions of an MCR problem, see [26]. Unimodality regions in the Hidden-AFS are calculated similarly to monotonicity, i.e., evaluating a grid with an appropriate criterion. Here the maximum in the profile is found and then, similarly to Sec. 2.5, it is checked whether the profile is monotone increasing before the maximum and monotone decreasing after the maximum. A similar approach can also be used for noisy data. However, useful statements about the geometric structure of unimodal profiles are not easy to make when compared to monotonicity. This is due to the fact that the convex combination of two unimodal profiles is not necessarily unimodal.

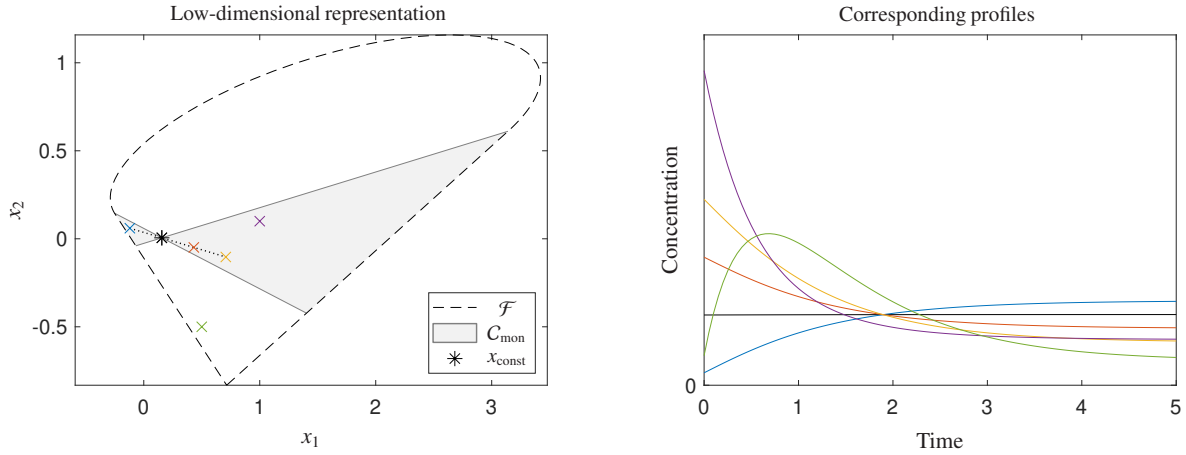

Figure 5: The low-dimensional representation of the model problem from Example 1.1 where the points in C_{mon} represent monotone profiles (left). The profiles corresponding to the marked points are shown on the right side. The black profile is constant and corresponds to x_{const}. As an example for Lemma 2.10, the blue and yellow profiles can be given as a linear combination of the black and the red profiles. Their low-dimensional representation is located on the dotted line segment. Also, the green point is not located in C_{mon} and the corresponding profile is not monotone.

Therefore, the structure seen in Fig. 9 of [29], see also the green point and profile in Fig. 5, cannot be generalized and is only seen for this particular kinetics. However, it would be possible to generally apply similar geometric considerations as in Sec. 3.3. This situation illustrates that mixed constraints can also be applied, i.e., monotonicity can be applied to some profiles and unimodality to others. Other meaningful constraints can also be used, for example, that the profiles must be approximated by an exponential function. In Sec. 3.4 we show an example of how mass balance can be used as a constraint.

3. Simulated examples and calculation methods

3.1. Michaelis-Menten kinetics - a rank-3 model problem

Next, a model data set with an underlying Michaelis-Menten kinetics is introduced. This model problem mirrors the behavior of the experimental data set in Sec. 4. A similar problem was presented in Sec. 6 of [25], but here it is slightly modified to provide an example with hidden rank deficiency. More information on modeling kinetics with rank deficiency is given in [18].

Example 3.1. *The factor C is computed by using the kinetics*

$$\boldsymbol{S} + \boldsymbol{E} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \boldsymbol{[E\text{-}S]} \xrightarrow{k_2} \boldsymbol{P} + \boldsymbol{E}$$

with $k_1 = 10$, $k_{-1} = 0.1$ *and* $k_2 = 3$ *and the initial concentrations* $c_S(0) = 1$, $c_E(0) = 0.1$ *and* $c_{[E-S]}(0) = c_P(0) = 0$. *The time interval* $[0, 7.5]$ *and a grid with* $k = 101$ *equidistant points are used. The pure component spectra are computed by using shifted Gaussians as in Example 1.1 with the parameters* $c_1 = 20$, $c_2 = 40$, $c_3 = 60$, $c_4 = 80$, $d_1 = d_2 = d_3 = 200$, $d_4 = 100$, $b_1 = 0.03$, $b_2 = 0.02$, $b_3 = 0.015$, $b_4 = 0.011$, $a_1 = a_4 = 1$, $a_2 = a_3 = 10$ *(where* $(1, 2, 3, 4) = (\boldsymbol{S}, \boldsymbol{E}, \boldsymbol{[E\text{-}S]}, \boldsymbol{P})$*). For the spectra the interval* $x \in [1, 100]$ *and* $n = 100$ *equidistant points are used. Furthermore the first 6 time steps of the factor C are discarded for the analysis. So it is* $D = C_{7:101,:}\, S^T$ *with* $D \in \mathbb{R}^{95\times 100}$.

After discarding the first 6 time steps, the model problem has only monotone concentration profiles, see Fig. 6. Furthermore, Fig. 7 demonstrates that this problem is rank-deficient, since $\text{rank}(D) = \text{rank}_+(D) = 3$ and the AFS for three components is not empty. However, only one low-dimensional representation of the original concentration profiles is located in the AFS. This is an indication of hidden rank deficiency. If we additionally apply the constraint of monotonicity, then $\text{rank}_f(D) = 4$ for this data set. Then the Hidden-AFS can be calculated and the ambiguity of the solutions can be approximated. The feasible bands from both approaches are compared in Fig. 8 and Fig. 9.

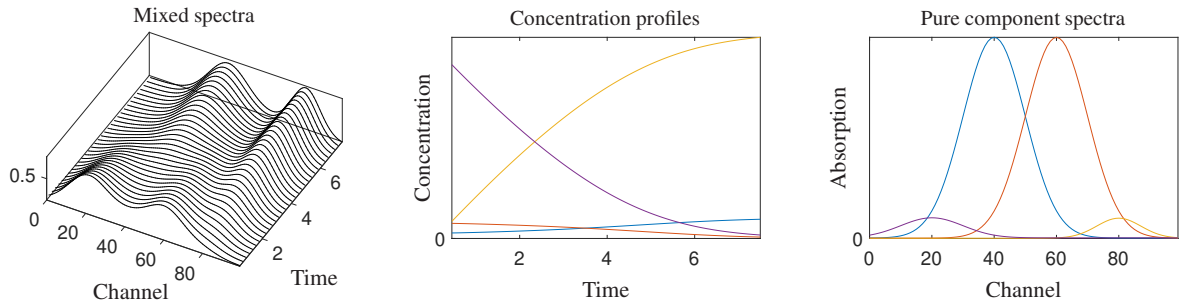

Figure 6: The data matrix and pure component factors for the model problem from Example 3.1. The first 6 time steps are discarded.

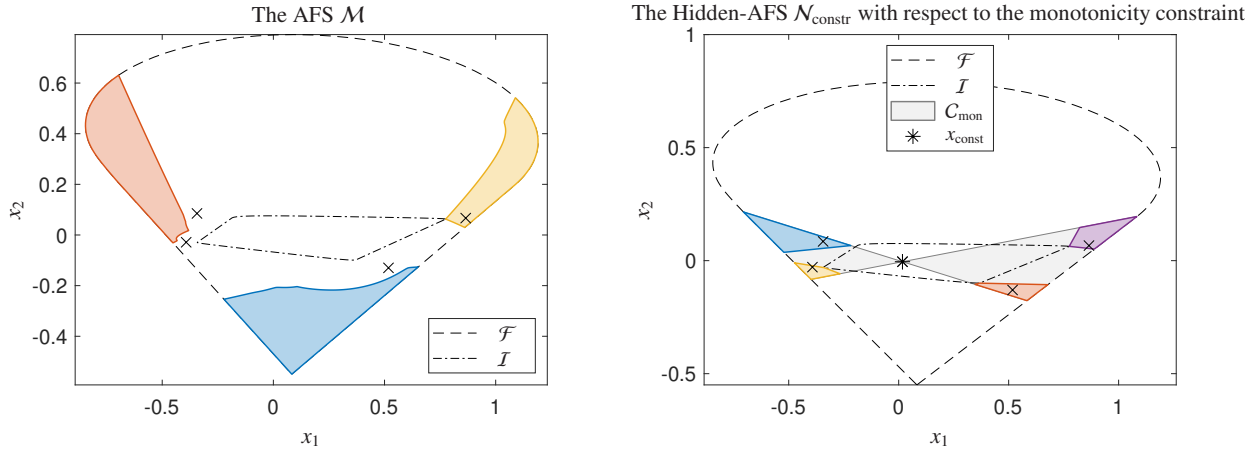

Figure 7: The AFS and the Hidden-AFS for the factor C for the Example 3.1. Left: The AFS show the ambiguity of the solutions with three chemical components. Only one of the points representing the true profiles ($\times$) is included in the AFS. Right: The Hidden-AFS with respect to the monotonicity constraint and rank deficiency. Each of the four isolated Hidden-AFS subsets represents the ambiguity of a monotone concentration profile of a pure component. The concentration profiles are chemically meaningful, $\text{rank}_f(D) = 4$ and no nested triangle with vertices in $\mathcal{N}_{\text{constr}}$ exists.

3.2. A numerical approach

The most straightforward method to compute the Hidden-AFS for problems with hidden rank deficiency is based on a numerical evaluation of points between $\mathcal{I}$ and $\mathcal{F}$. Here a method like the polygon inflation algorithm [24] proves to be useful. Such methods help to approximate the Hidden-AFS by evaluating a function for different points and deciding, whether they are located in the Hidden-AFS. This decision is mainly driven by the numerical solution of an optimization problem.

The objective function was modified to account for problems with open rank deficiency in Sec. 5.1 of [25]. Since the modifications are essentially the same in the case of hidden rank deficiency, they will not be repeated in full here. The main difference from the calculation for rank-regular problems, is the additional computational cost of calculating the factor, which cannot be expressed with the help singular vectors of the data matrix. An NNLS solver can be used for this, see, e.g., [14].

Problems with hidden rank deficiency require a further modification of the objective function, namely, the use of a constraint function. This is described in Sec. 2.5 and Sec. 2.6 and more details can be found in the literature, in particular, in [26] the implementation of the constraints for the objective function is considered. The evaluation of the constraint function further increases the computation time for problems with hidden rank deficiency. Thus, the main drawback of the numerical approach is the comparably large computation time due to the optimization of the objective function. However, it is suitable for experimental and noisy data sets.

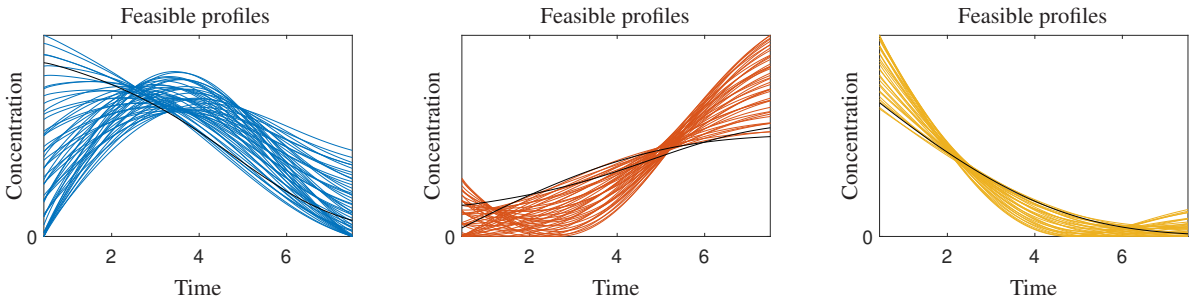

Figure 8: The feasible bands for the model problem from Example 3.1 without any additional knowledge on the solutions. Since $\operatorname{rank}(D) = \operatorname{rank}_+(D) = 3$ the AFS is not empty. The true profiles are plotted as black lines and they are close to the feasible bands, but only the feasible band on the right includes the original solution. The feasible band contains many solutions that are not monotone, i.e., not chemically meaningful in this case.

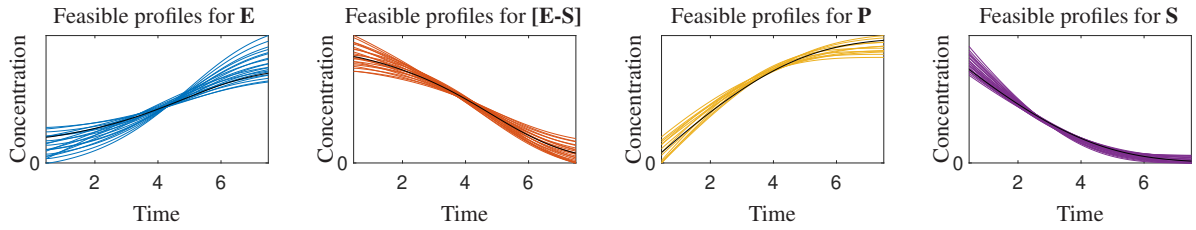

Figure 9: The feasible bands for the model problem from Example 3.1 with respect to the monotonicity constraint and rank deficiency. Unlike the solutions in Fig. 2, the true profiles (black lines) are included in the feasible bands. The feasible bands only contain monotone profiles here.

3.3. A geometric approach

For some cases, a very simple and fast geometric solution can be provided. This is similar to the concept of Borgen triangles in [20] or the calculation of the AFS when one or more components are known. The main advantage of the geometric approach is the very low computational cost. This simple geometric approach to solving a problem with hidden rank deficiency is described for the problem from Example 3.1, i.e., a 4-component case with monotonicity constraint, where two components are increasing and two components are decreasing.

Since the monotonicity boundaries (i.e. the boundaries of C_{mon}) are given by two lines, see Lemma 2.12 for $s = 3$, there are four points, where these lines cross $\mathcal{F}$. We call these points the monotonicity boundary points. An additional $\mathcal{I}$-monotonicity condition must be satisfied: the quadrilateral spanning the four monotonicity boundary points on $\mathcal{F}$ has to include $\mathcal{I}$. Two problematic situations are possible:

1. $\mathcal{I}$ intersects the quadrilateral of the four monotonicity boundary points in the non-monotone part (green lines in Fig. 10). Then no solutions to the constrained problem are possible, since some of the data points of $\mathcal{I}$ cannot be described by feasible monotone profiles, see also [22]. In other words, there exists no nested quadrilateral with monotonicity constraint.
2. $\mathcal{I}$ intersects the quadrilateral in the monotone part (blue lines in Fig. 10, note that the blue lines must not always overlap with $\mathcal{F}$). Then the geometric approach should be adapted and is more complex than explained here.

See Fig. 10 for illustrates a situation where none of the problematic situations occur and the $\mathcal{I}$-monotonicity condition is satisfied.

If the conditions are met, then the Hidden-AFS is approximated by finding tangents from each of the four monotonicity boundary points on $\mathcal{F}$ to $\mathcal{I}$, see Fig. 10 (right). The tangent that starts opposite to the currently examined subset does not directly affect the subset of the Hidden-AFS. However, from the point where it crosses $\mathcal{F}$, another tangent to $\mathcal{I}$ is constructed and part of it encloses the isolated subset of the Hidden-AFS. This means, that the Hidden-AFS is enclosed by $\mathcal{F}$, the boundary of the monotonicity constraint, and two tangents. See Fig. 10 for an illustration.

Due to the fact that the intersection of $\mathcal{F}$ with the area of monotonicity constraint is usually not a convex polygon (see Remark 2.11), the classic rotation algorithm (see [5] or [23] for an application to problems with open rank

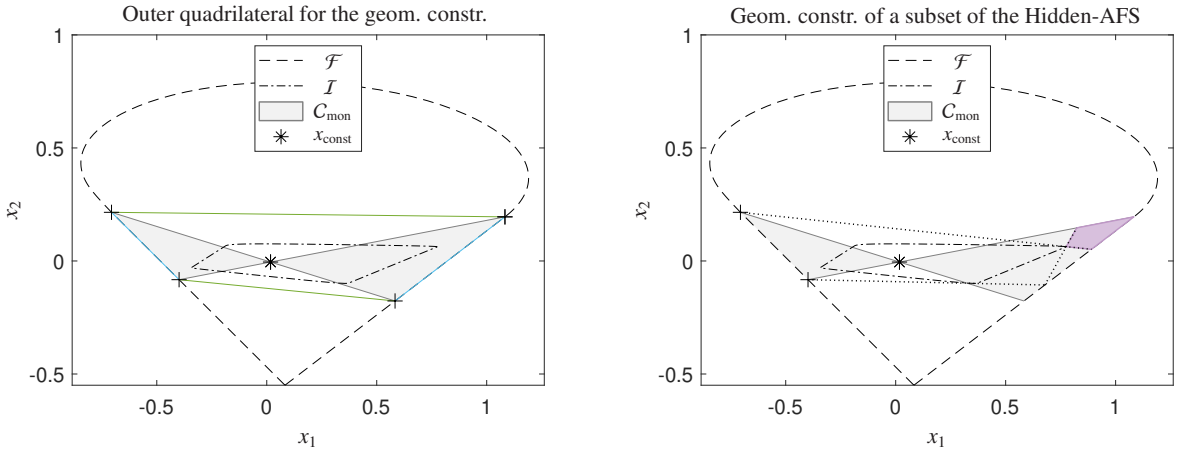

Figure 10: $\mathcal{I}$ and $\mathcal{F}$ for the factor C for the Example 3.1. Left: The monotonicity boundary points are marked with a +. The green and blue lines do not intersect $\mathcal{I}$ and satisfy a condition for the geometric construction. Right: The geometric construction of one isolated subset of the Hidden-AFS (purple) is shown. Dotted lines represent tangents to $\mathcal{I}$, which are used to construct this subset of the Hidden-AFS. A black + denotes a monotonicity boundary point where the construction for this subset of the Hidden-AFS starts.

deficiency) is not applicable here. When rotating the polygon, the goal is to move the final point as close to the origin as possible (for a given start point or edge). However, due to the lack of convexity, the tangents of $\mathcal{I}$ do not always lead to the innermost point.

Therefore, the idea of the geometric approach is to start with the largest possible polygon spanning the four monotonicity boundary points and then bring each of the four vertices as close to the origin as possible. This is achieved with the help of tangents.

It is possible to adapt this approach for various different cases, but a simple rotation algorithm covering all possible cases of the hidden rank deficiency in a similar way to problems with open rank deficiency ([23]) does not seem to be possible. This is due to the fact that the intersection of $\mathcal{F}$ with the area of constraints is not necessarily convex, nor does it contain $\mathcal{I}$.

3.4. Using mass balance as a constraint

Rank deficiency in spectroscopic data sets is often caused by multiple mass balance constraints. For example, there are two mass balance constraints in the Michaelis-Menten data set from Example 3.1:

$$\begin{aligned} c_S + c_P + c_{[E-S]} &= \text{const} \\ c_E + c_{[E-S]} &= \text{const} \end{aligned}$$

where c_S, c_E, c_P and $c_{[E-S]}$ denote the concentration profiles of the pure components. These relationships can be further used to reduce the ambiguity of the solutions, similarly to Lemma 2.10 and Remark 2.11. If the convex combination of two profiles results in a constant profile, then the low-dimensional representations of the corresponding profiles are located on a line passing through x_{const}. If three profiles satisfy a mass balance constraint, then x_{const} is located in the triangle that is spanned between them.

Let us consider the mass balance $c_E + c_{[E-S]} = \text{const}$, which states that a linear combination of the concentration profiles of **E** and **[E-S]** results in a constant profile. According to Lemma 2.8, this constant profile has a low-dimensional representation x_{const}. This constant profile is scaled if and only if the linear combination of two scaled profiles is a convex combination, see the reasoning in the proof of Lemma 2.10 (negative coefficients are not allowed). Then according to Lemma 2.10 x_{const} is located on a line between the low-dimensional representations of **E** and **[E-S]**. Conversely this means, that if the concentration profile of **E** is given, then the Hidden-AFS subset of **[E-S]** is reduced to a line, because x_{const} is fixed. See Fig. 11 for an illustration of this effect.

Let us consider the other mass balance $c_S + c_P + c_{[E-S]} = \text{const}$. The statement of Lemma 2.8 is analogously extended to triangles. The constant profile is scaled if and only if the linear combination of the three scaled profiles is

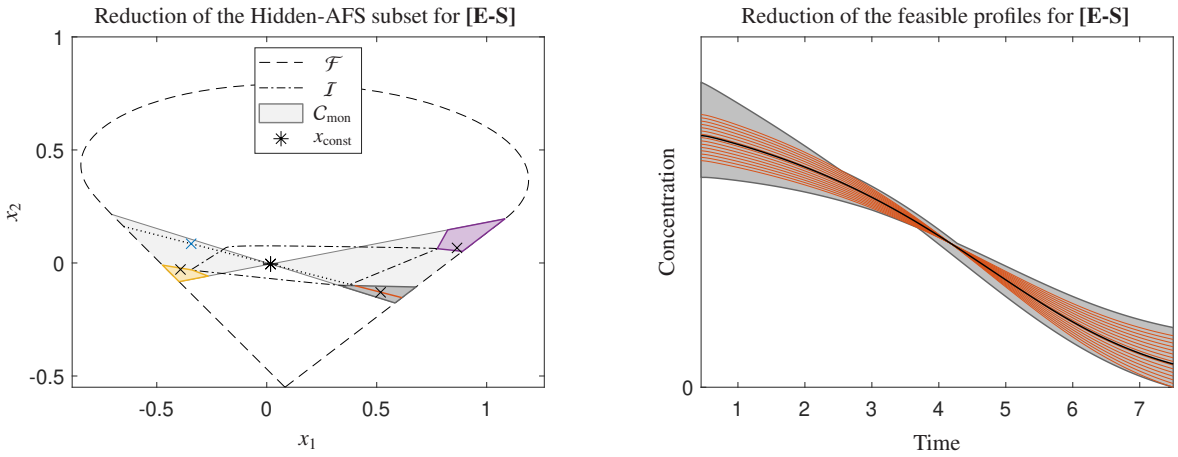

Figure 11: The Hidden-AFS for the factor C for Example 3.1, with the mass balance constraint $c_E + c_{[E-S]} = \text{const}$. Left: The low-dimensional representations of the true profiles are shown as black ×. The concentration profile of **E** is assumed to be given and is shown with as blue ×. The isolated Hidden-AFS subset for **E** is not shown. The original Hidden-AFS subset of **[E-S]** is shown in dark gray and the reduced subset is shown in its original orange color. Auxiliary lines for the geometric construction are dotted. The purple and the yellow Hidden-AFS subsets are unchanged. Right: The original feasible profiles (gray band) and the reduced feasible profiles (orange).

a convex combination. Then the low-dimensional representation of a constant profile x_{const} is the convex combination of the the low-dimensional representations of the concentration profiles of **S**, **P** and **[E-S]**. According to Fig. 12, if the concentration profile of **S** is given, then the Hidden-AFS subset **P** is reduced, because the triangle spanning the three low-dimensional representations must include x_{const}.

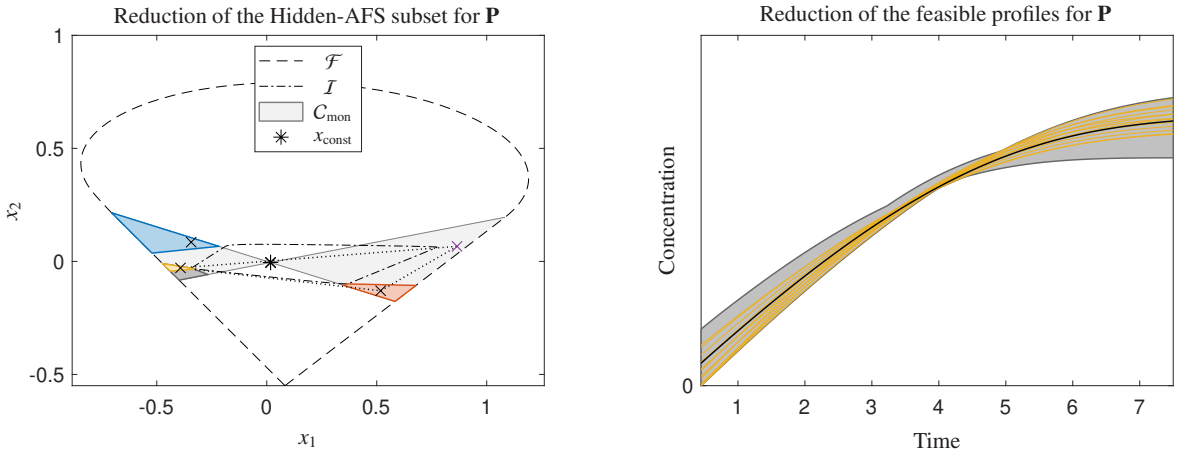

Figure 12: The Hidden-AFS for the factor C for Example 3.1, with the mass balance constraint $c_S + c_P + c_{[E-S]} = \text{const}$. Left: The low-dimensional representations of the true profiles are shown as ×. The concentration profile of **S** is assumed to be given and is shown with as purple ×. The Hidden-AFS subset for **S** is not shown. The original Hidden-AFS subset of **P** is shown in dark gray and the reduced subset is shown in the original yellow color. Auxiliary lines for the geometric construction are dotted. The red and the blue Hidden-AFS subsets are unchanged. Right: The original feasible profiles (gray band) and the reduced feasible profiles (yellow).

Depending on the situation, the use of the mass balance constraint can provide a lot of information during pure component decomposition. However, this approach is sensitive to noise. Even a small error in the location of x_{const} can lead to incorrect conclusions. Therefore, the mass balance constraint as described here should only be applied after a careful consideration.

3.5. Rank-2 problems

Systems of rank 2 are an exception to the proposed approach. According to the proof in [28], a rank 2 matrix always has the nonnegative rank $m = 2$. Therefore, no open rank deficiency is possible for systems with rank $s = 2$. Furthermore, no additional constraints can help to resolve this case, as seen in the previous examples, because no additional constraints can transform the hidden rank deficiency into an open one. However, they can still help to reduce the ambiguity of the solutions. We show this with the help of an example.

Example 3.2. *The factor $C \in \mathbb{R}^{201\times 3}$ is computed for the kinetics*

$$\boldsymbol{A} + 2\boldsymbol{B} \rightarrow \boldsymbol{C}$$

with the kinetic constant $k = 1$ and the initial concentration $C_{1,:} = (1, 0.4, 0)$ on a grid of $k = 201$ equidistant points. The factor S is defined similarly to Example 1.1 with the parameters $a_1 = 2.5$, $a_2 = 25$, $a_3 = 6$, $c_1 = 20$, $c_2 = 40$, $c_3 = 60$, $d_1 = d_2 = d_3 = 200$, $b_1 = b_2 = 0.075$, $b_3 = 0.065$ on an equidistant grid with $n = 100$ points. The profiles in Fig. 13 and Fig. 14 use a different scaling. This data set is rank-deficient, because the kinetics lead to two mass balance constraints.

Geometrically, $\mathcal{I}, \mathcal{F}$ and $\mathcal{M}$ are located in a 1-dimensional space, i.e., they are on a line. One of the pure component spectra is located to the left and the other to the right of $\mathcal{I}$. The same is true for concentration profiles. If there are more components, then according to Remark 2.6 their corresponding low-dimensional representations can be located anywhere in $\mathcal{F}$, see Fig. 14. However, at least two points should be chosen according to the AFS for two components. Therefore, this problem should firstly be solved as a rank-regular problem and the remaining profiles should be calculated afterwards.

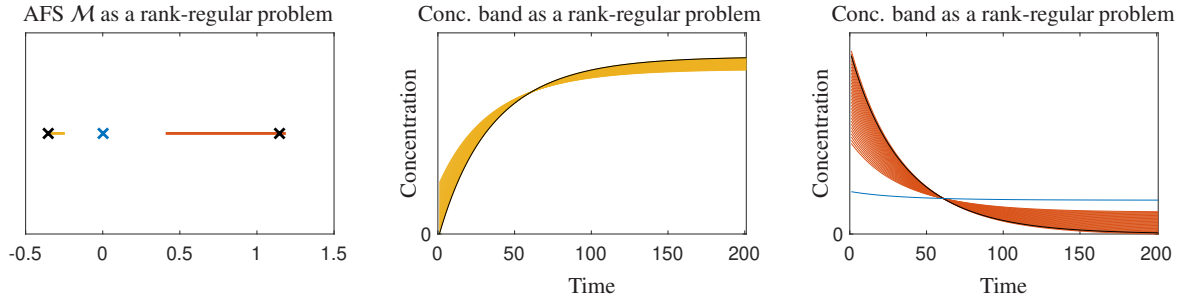

Figure 13: The one-dimensional AFS (left) for the model data set from Example 3.2 as a rank-regular problem. The feasible bands (center and right) are colored the same as the corresponding sections of the AFS. The original concentration profiles are shown in black and blue color. The blue original concentration profile is outside of the feasible bands of a rank-regular problem.

3.6. Trivial rank deficiency: constant concentrations or multiple products

Sometimes data sets with hidden rank deficiency cannot be resolved even when constraints are applied. This can happen, when the condition $\operatorname{rank}_f(D) = r$ in Def. 2.5 does not hold. Then a feasible solution with fewer components and chemically meaningful pure component concentration profiles can be possible. We demonstrate this in two cases: when a pure component has a constant concentration profile and when a reaction results in multiple products.

First, let us consider the reaction

$$\mathbf{A} \rightarrow \mathbf{B} \rightarrow \mathbf{C} + \mathbf{D}.$$

In this case, we observe rank deficiency because the concentration profiles of **C** and **D** are identical. However, they have exactly the same low-dimensional representation. This means, that the low-dimensional representation is a quadrilateral with two overlapping vertices. In other words, we can consider this as a rank-regular three-component problem with **A**, **B** and **[C-D]**. The spectrum of **[C-D]** is the sum of the pure component spectra of **C** and **D**.

Second, let us consider the reaction

$$\mathbf{A} \rightarrow \mathbf{B} \rightarrow \mathbf{C}$$

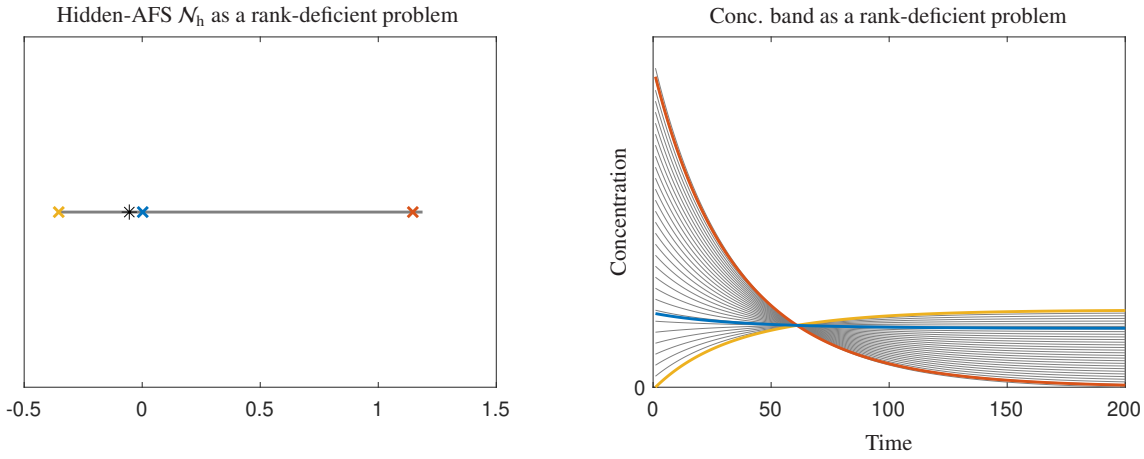

Figure 14: The one-dimensional Hidden-AFS (left) for the model data set from Example 3.2, if the rank deficiency is considered, but no further constraints are applied. The Hidden-AFS overlaps with $\mathcal{F}$ and the concentration profiles are located in the band of feasible concentration profiles (right). The colors in the Hidden-AFS correspond to the concentration profiles. Black asterisk shows the location of x_{const}.

with the time-independent component, e.g., solvent,

D

which has a constant concentration profile with the low-dimensional representation x_{const}. Similar to the previous problem, this is also a case of rank deficiency because there are two mass balance constraints. Furthermore, this problem has a hidden rank deficiency due to Remark 2.11. Such an example was also given in Sec. 4.2 in [22] for the alkaline hydrolysis process. This problem can also be considered as a rank-regular three-component problem, but the correct pure component spectra cannot be directly seen in the solution. A possible approach would be to add x_{const} to the constructed triangle and then calculate the band boundaries for the spectral profiles, see [22].

In both of these cases, it is not possible to detect hidden rank deficiency only with the information in the data set. The calculated concentration profiles of the rank-regular problem are chemically meaningful. Furthermore, the low-dimensional representation of the profile that is missing in the rank-regular solution is located in or on the triangle representing the rank-regular solution. Nevertheless, Remark 2.6 can be applied as shown in the two examples.

These are just a few examples of situations where constraint functions do not help to resolve hidden rank deficiency. The only way to solve such problems seems to be the direct use of the additional information in the low-dimensional representation of the data set.

4. Results for an experimental data set

4.1. Rhodium catalyzed hydroformylation data set

In this section a data set based on the in-situ spectroscopic monitoring of the rhodium catalyzed hydroformylation of 3,3-dimethly-1-butene with a rhodium/tri(2,4-di-*tert*-butylphenyl)phosphite catalyst in *n*-hexane is analyzed. The catalytic experiment was conducted at [Rh] = 0.3 mM, [TDTBPP] = 6 mM, [alkene] = 0.45 M at a temperature of 30 °C, p(H_2) = 10 bar and p(CO) = 2 bar. The data set consists of $k = 300$ spectra and two intervals of spectral channels are analyzed. The first window contains wave numbers in $[1964, 2107]\,\text{cm}^{-1}$ and in this window the olefin, the acyl complex and the hydrido complex show IR absorption bands. The second window contains wave numbers in $[3411, 3479]\,\text{cm}^{-1}$ and the contribution to an absorption band from the aldehyde. The combination of both intervals results in $n = 879$ spectral channels.

Since the second interval contains the absorption band of only one component, the corresponding concentration profile can be extracted directly. However, in order to demonstrate the approach for a problem with hidden rank deficiency, both intervals are combined for the analysis, see Fig. 15. Further information on this system is given in [13]. The true profiles and the first 30 singular values of this data set are also shown in Fig. 2 in [18].

The concentration profiles of the four components are determined by a Michaelis-Menten kinetics, similar to Example 3.1. Therefore, this is a rank-deficient problem. Furthermore, this problem suffers from hidden rank deficiency because a further analysis reveals that $\mathrm{rank}_+(D) = 3$ with respect to perturbations, see Fig. 16. The results are verified by a kinetic model. The low-dimensional representations of the concentration profiles are compared with the Hidden-AFS and the comparison in feasible bands is shown in Fig. 17.

4.2. Approach to noisy data

Noisy data pose several challenges to the computation of the Hidden-AFS. Typically, numerical methods are preferred because they can better account for perturbations and negativity in the data set.

The baseline of the hydroformylation data set is corrected and the data matrix contains only small negative entries. However, the peaks in the FTIR spectroscopy are sharp and the absorption for several wave numbers is close to zero for all spectra. Therefore the perturbations in D have a particularly large effect on the data points a_i for $i = 1 \ldots, 879$, which form $\mathcal{I}$ in the noise-free case. For example, several points are not even included in $\mathcal{F}$, see Fig. 15.

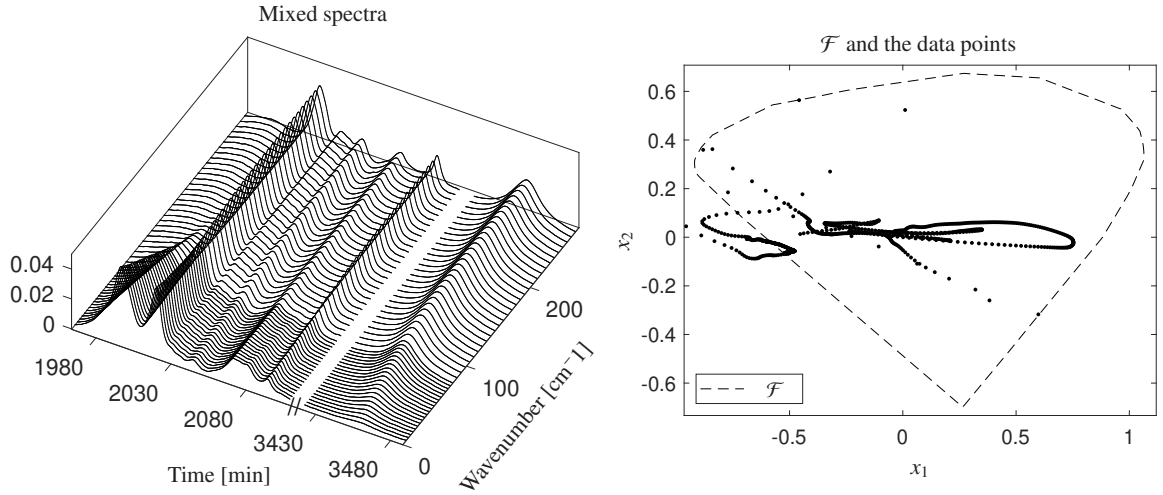

Figure 15: Left: the hydroformylation data set from Sec. 4.1. Right: The outer boundary $\mathcal{F}$ and the data representing points for this data set. Some data points are outside of this illustration. The most distant data point is located at approximately (-9.8,4.9).

To compute the Hidden-AFS despite these problems, a control parameter is used to accept relative slightly negative entries in both factors. This control parameter was set to $\varepsilon = 0.012$, see [24] for more details on this approach.

Monotonicity constraints are satisfied in this data set, because the first spectrum was taken at 4.94 min. Another control parameter is used to deal with perturbations. Thus, small deviations from monotonicity are also accepted. The control parameter for accepting slightly non-monotone profiles was set to $\rho = 0.015$, see [26] for more details. Also the low-dimensional representation of a constant profile x_{const} was found as a least squares approximation due to perturbations.

4.3. The Hidden-AFS for the data set

The AFS for the data set is computed by the polygon inflation algorithm. The AFS for three components is shown along with the low-dimensional representation of the true profiles in the left plot of Fig. 16. The computation time (on a standard PC with a 3.4 GHz CPU and 16 GB RAM; using only one core) with the polygon inflation method was about 5.6 seconds. The true profiles were computed using hard modeling with a kinetic model, see [13]. However, only one of the four points is included in the AFS, since the constrained rank of the data matrix is $\mathrm{rank}_f(D) = 4$. This indicates that the data set suffers from hidden rank deficiency and suitable constraints should be used.

Soft-constraints for monotonicity in factor C are used in order to compute the Hidden-AFS with respect to hidden rank deficiency. Then only chemically meaningful solutions (in this case, with monotone profiles) are represented with the Hidden-AFS. The four isolated subsets are shown in the right plot of Fig. 16. They are located in the area

of monotone profiles C_{mon}. They are found with the polygon inflation method and the new cost function for rank-deficient problems, see Sec. 3.2. The four low-dimensional representations of the true solution are included in the Hidden-AFS. The feasible concentration profiles for the four components are shown in Fig. 17.

The computation time of the Hidden-AFS on the same PC was about 72.4 seconds. The large difference in the computation time when compared to the AFS is caused by the different objective function to classify a point as feasible or not. For problems with rank deficiency, this includes the approximation of a factor that has a higher rank than the data matrix D, see Sec. 3.2.

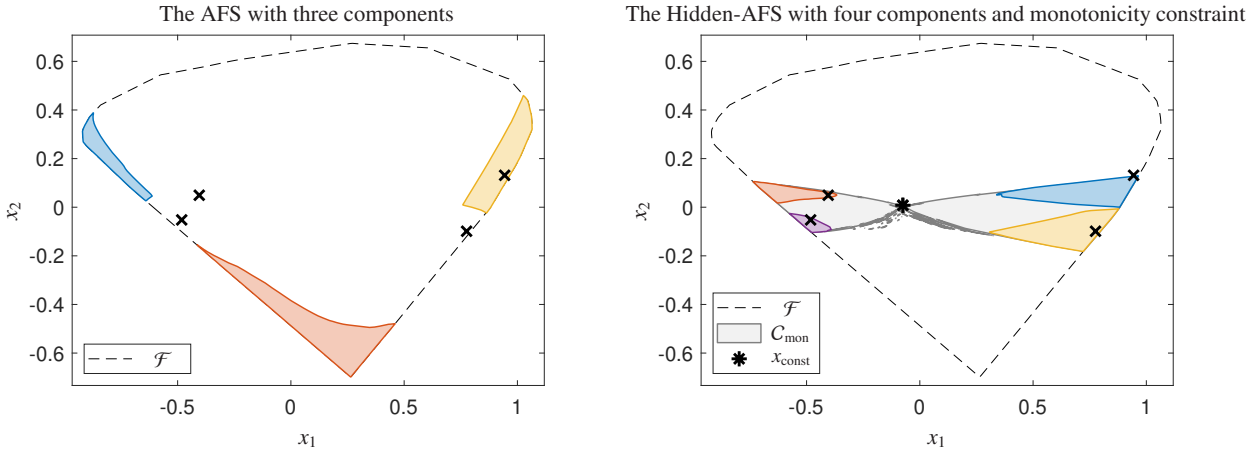

Figure 16: The AFS and the Hidden-AFS for factor C of the hydroformylation data set from Sec. 4.1. The true solution from a kinetic model is shown as ×. Left: The AFS consists of three separate isolated subsets, one of which does not contain any monotone concentration profiles. This indicates hidden rank deficiency. Also, only one of the points from the true solution (×) is included in the AFS. Right: The Hidden-AFS with respect to the monotonicity constraint (C_{mon}) and hidden rank deficiency. Each of the four isolated subsets contains one point of the true solution.

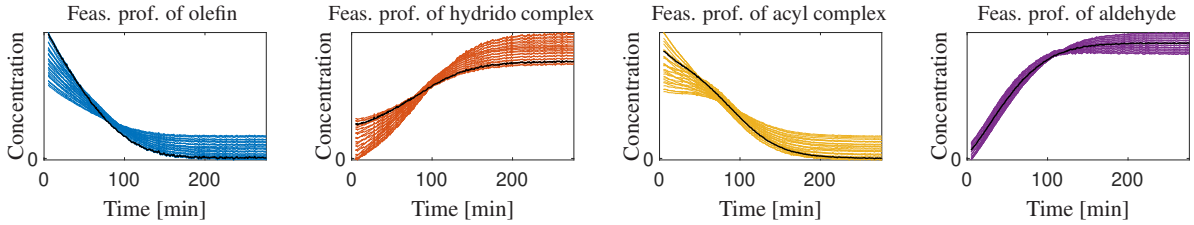

Figure 17: The feasible bands for the hydroformylation data set from Sec. 4.1 with respect to hidden rank deficiency and the monotonicity constraint. The true profiles (black lines) are included. The first spectrum was taken at 4.94 min.

5. Conclusion

Detecting a hidden rank deficiency is like finding the true reality one level below a feigned and simplified reality. This is similar to the shadows on the cave walls in Plato's allegory, where the shadows are only projections of the true reality. Like the projections on the cave walls, the true nature of the hidden rank-deficient problems is obscured because it is located in a higher dimension. For a given data set, it is not clear a priori where the true reality can be found. However, this paper demonstrates that the problem can be defined and understood.

Admittedly, it is a rare case where a sound chemical interpretation is possible for a data set, but the true chemical interpretation lies deeper. In any case, additional constraints, such as monotonicity or mass balance, can help to resolve a hidden rank deficiency and determine the ambiguity of the true solution. We note that while the geometric approach is simple and fast, it is not always feasible. When geometric methods are not applicable, more robust numerical methods can be used. This is the case for noisy experimental data. When rank-2 problems suffer from a

rank deficiency, it is always a hidden rank deficiency. This case, as well as some other cases where rank deficiency appears trivially, are particularly challenging to resolve.

References

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

[2] M. Amrhein, B. Srinivasan, D. Bonvin, and M. M. Schumacher. On the rank deficiency and rank augmentation of the spectral measurement matrix. *Chemom. Intell. Lab. Syst.*, 33(1):17–33, 1996.

[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] J. E. Cohen and U. G. Rothblum. Nonnegative ranks, decompositions, and factorizations of nonnegative matrices. *Linear Algebra Appl.*, 190:149–168, 1993.

[7] A. de Juan, Y. Vander Heyden, R. Tauler, and D. L . Massart. Assessment of new constraints applied to the alternating least squares method. *Anal. Chim. Acta*, 346(3):307–318, 1997.

[8] P.J. Gemperline and E. Cash. Advantages of soft versus hard constraints in self-modeling curve resolution problems. Alternating least squares with penalty functions. *Anal. Chem.*, 75:4236–4243, 2003.

[9] N. Gillis. *Nonnegative Matrix Factorization*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2020.

[10] N. Gillis and F. Glineur. On the geometric interpretation of the nonnegative rank. *Linear Algebra Appl.*, 437(11):2685 – 2712, 2012.

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

[12] G. H. Golub and C. F. Van Loan. *Matrix computations*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, 2013.

[13] C. Kubis, M. Sawall, A. Block, K. Neymeyr, R. Ludwig, A. Börner, and D. Selent. An operando FTIR spectroscopic and kinetic study of carbon monoxide pressure influence on rhodium-catalyzed olefin hydroformylation. *Chem.-Eur. J.*, 20(37):11921–11931, 2014.

[14] C. L. Lawson and R. J. Hanson. *Solving least squares problems*, volume 15 of *Classics Appl. Math.* Society of Industrial and Applied Mathematics, Philadelphia, 1995.

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

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

[17] K. Neymeyr and M. Sawall. On the set of solutions of the nonnegative matrix factorization problem. *SIAM J. Matrix Anal. Appl.*, 39:1049–1069, 2018.

[18] K. Neymeyr, M. Sawall, and T. Andersons. A note on rank deficiency and numerical modeling. *J. Chemom.*, page e3550, 2024.

[19] N. Omidikia, S. Beyramysoltan, J. Mohammad Jafari, E. Tavakkoli, M. Akbari Lakeh, M. Alinaghi, M. Ghaffari, S. Khodadadi Karimvand, R. Rajkó, and H. Abdollahi. Closure constraint in multivariate curve resolution. *J. Chemom.*, 32(12):e2975, 2018.

[20] 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.

[21] J. Saurina, S. Hernández-Cassou, R. Tauler, and A. Izquierdo-Ridorsa. Multivariate resolution of rank-deficient spectrophotometric data from first-order kinetic decomposition reactions. *J. Chemom.*, 12(3):183–203, 1998.

[22] M. Sawall, T. Andersons, H. Abdollahi, S. K. Karimvand, B. Hemmateenejad, and K. Neymeyr. Calculation of lower and upper band boundaries for the feasible solutions of rank-deficient multivariate curve resolution problems. *Chemom. Intell. Lab. Syst.*, 226:104577, 2022.

[23] M. Sawall, T. Andersons, and K. Neymeyr. On the area of feasible solutions for rank-deficient problems: II. The geometric construction. *Chemom. Intell. Lab. Syst.*, 235:104782, 2023.

[24] 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.

[25] M. Sawall and K. Neymeyr. On the area of feasible solutions for rank-deficient problems: I. Introduction of a generalized concept. *J. Chemom.*, 35(3):e3316, 2020. e3316 cem.3316.

[26] M. Sawall, N. Rahimdoust, C. Kubis, H. Schröder, D. Selent, D. Hess, H. Abdollahi, R. Franke, Börner A., and K. Neymeyr. Soft constraints for reducing the intrinsic rotational ambiguity of the area of feasible solutions. *Chemom. Intell. Lab. Syst.*, 149, Part A:140–150, 2015.

[27] R. Tauler, A. Smilde, and B. Kowalski. Selectivity, local rank, three-way data analysis and ambiguity in multivariate curve resolution. *J. Chemom.*, 9(1):31–58, 1995.

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

[29] S. Vali Zade, M. Sawall, K. Neymeyr, and H. Abdollahi. Introducing the monotonicity constraint as an effective chemistry-based condition in self-modeling curve resolution. *Chemom. Intell. Lab. Syst.*, 190:22–32, 2019.