SIAM J. MATRIX ANAL. APPL.
Vol. 44, No. 1, pp. 27-52

PARTIAL IDENTIFIABILITY FOR NONNEGATIVE MATRIX FACTORIZATION*

NICOLAS GILLIS† AND RÓBERT RAJKÓ‡

Abstract. Given a nonnegative matrix factorization, R, and a factorization rank, r, exact nonnegative matrix factorization (exact NMF) decomposes R as the product of two nonnegative matrices, C and S with r columns, such as $R = CS^\top$. A central research topic in the literature is the conditions under which such a decomposition is unique/identifiable up to trivial ambiguities. In this paper, we focus on partial identifiability, that is, the uniqueness of a subset of columns of C and S. We start our investigations with the data-based uniqueness (DBU) theorem from the chemometrics literature. The DBU theorem analyzes all feasible solutions of exact NMF and relies on sparsity conditions on C and S. We provide a mathematically rigorous theorem of a recently published restricted version of the DBU theorem, relying only on simple sparsity and algebraic conditions: it applies to a particular solution of exact NMF (as opposed to all feasible solutions) and allows us to guarantee the partial uniqueness of a single column of C or S. Second, based on a geometric interpretation of the restricted DBU theorem, we obtain a new partial identifiability result. This geometric interpretation also leads us to another partial identifiability result in the case $r = 3$. Third, we show how partial identifiability results can be used sequentially to guarantee the identifiability of more columns of C and S. We illustrate these results on several examples, including one from the chemometrics literature.

Key words. nonnegative matrix factorization, uniqueness, identifiability, multivariate curve resolution, window factor analysis, self-modeling curve resolution

MSC codes. 15A23

DOI. 10.1137/22M1507553

1. Introduction.

Given a nonnegative matrix $R \in \mathbb{R}^{m\times n}_+$ and a factorization rank r, nonnegative matrix factorization (NMF) requires computing two nonnegative matrices, $C \in \mathbb{R}^{m\times r}_+$ and $S \in \mathbb{R}^{n\times r}_+$, such that $CS^\top \approx R$. NMF has become a standard technique in unsupervised data analysis and has found numerous applications, e.g., in hyperspectral imaging, audio source separation, topic modeling, and community detection, to cite a few; see, e.g., the books [8, 12] and the references therein. An application where NMF has been particularly popular is multivariate curve resolution (MCR) and self-modeling curve resolution, where the input matrix R represents the total response values from some chemical measurements of mixed samples. An example is when we consider the evolution of the spectral profile of a chemical reaction over time. More precisely, the ith row of R is the cumulative spectral content of the chemical reaction at the ith time step. An NMF of R, with $R(i,:) \approx C(i,:)S^\top$ for all i, provides the spectral signature of the chemical compounds in C, along with their proportion in the reaction over time in S. In general, the

* Received by the editors July 6, 2022; accepted for publication (in revised form) October 12, 2022; published electronically January 25, 2023.
https://doi.org/10.1137/22M1507553
Funding: The work of the first author was supported by the Fonds de la Recherche Scientifique-FNRS (F.R.S.-FNRS) and the Fonds Wetenschappelijk Onderzoek - Vlanderen (FWO) under EOS project O005318F-RG47, and by the Francqui Foundation.

† Department of Mathematics and Operational Research, University of Mons, 7000 Mons, Belgium (nicolas.gillis@umons.ac.be).

‡ Institute of Mathematics and Informatics, Faculty of Sciences, University of Pécs, Ifjúság u. 6. Pécs, H-7624, Hungary (rajko@gamma.ttk.pte.hu).

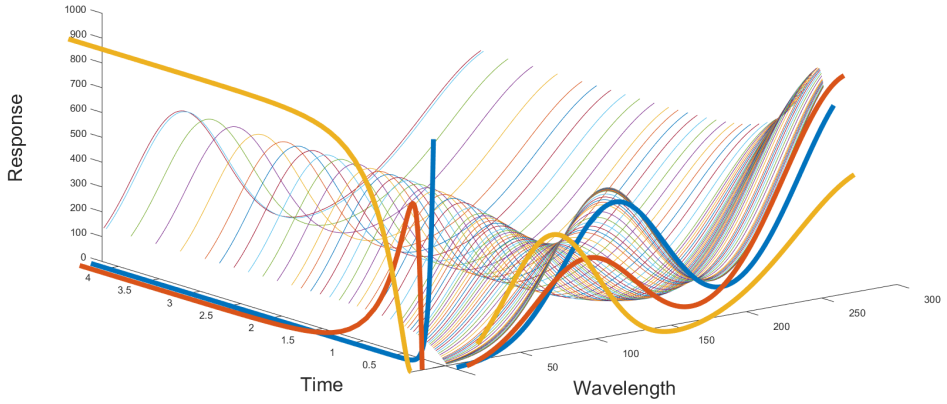

FIG. 1. *A three-component consecutive reaction example with the original composition in time and the signal in wavelength profiles.*

matrix C can be considered as the composition profile-matrix (each column of matrix C is a composition profile of a chemical, e.g., in a reaction in time), and the matrix S is the signal profile-matrix (each column of S will be the spectrum of a chemical). This model can cover most types of nonnegative measurement matrices and has been used successfully in chemistry, physics, biology, engineering, and informatics [7, 8, 9, 12, 24, 27, 41]. We provide a consecutive reaction example in which the reactant X forms an intermediate Y and the intermediate forms the product Z in two irreversible first-order reactions: $X \overset{k_1=20}{\longrightarrow} Y \overset{k_2=3}{\longrightarrow} Z$, where k_1 and k_2 are the first and the second reaction rate constants, respectively. Figure 1 depicts the data matrix curves and the original composition and signal profiles for the three components, X in "navy blue," Y in "chocolate," and Z in "gold tips" colors. See section 5.3 for another example.

Uniqueness/Identifiability. A crucial question in many applications is the uniqueness of a decomposition $CS^\top$ up to permutation and scaling, which is also known as the identifiability of $CS^\top$. In fact, uniqueness/identifiability (we will use both words interchangeably without attributes) allow NMF to recover the groundtruth factors that generated the data, such as the sources in audio source separation, the materials in hyperspectral images, and the chemical components in a reaction; see the discussions in [11] and [12, Chapter 4] and the references therein. To attack this question, we focus in this paper on exact NMF (that is, an errorless reconstruction), defined as follows.

DEFINITION 1 (exact NMF of size r). *Given a nonnegative matrix $R \in \mathbb{R}^{m\times n}$, the decomposition $CS^\top$, where $C \in \mathbb{R}^{m\times r}_+$ and $S \in \mathbb{R}^{n\times r}_+$, is an exact NMF of R of size r if $R = CS^\top$.*

Let us formally define the full uniqueness/identifiability of an exact NMF.

DEFINITION 2 (full identifiability of exact NMF). *The exact NMF of $R = C_\star S_\star^\top$ of size r is (fully) identifiable (also known as (a.k.a.) unique or essentially unique) if and only if, for any other exact NMF of $R = CS^\top$ of size r, there exist a permutation matrix $\Pi \in \{0,1\}^{r\times r}$ and a nonsingular diagonal scaling matrix D such that*

$$C = C_\star \Pi D \quad \text{and} \quad S^\top = D^{-1}\Pi^\top S_\star^\top.$$

In other words, any other exact NMF of $R = CS^\top$ of size r has the form

$$CS^\top = \sum_{k=1}^{r} C(:,k)S(:,k)^\top = \sum_{k=1}^{r} \underbrace{\alpha_k C_\star(:,\pi_k)}_{C(:,k)} \underbrace{\frac{1}{\alpha_k} S_\star(:,\pi_k)^\top}_{S(:,k)^\top} \tag{1}$$

for some permutation π of $\{1,2,\ldots,r\}$ and some positive scalars α_k $(1 \le k \le r)$.

In the NMF literature, all works we are aware of have focused on the full identifiability of exact NMF, and actually this is simply referred to as the identifiability of exact NMF. The chemometrics literature has been interested in the question of partial identifiability: when all the chemical components are not identifiable, it asks whether a subset of the profiles of these chemical components is identifiable. In the chemometrics literature that studies the MCR problem, the following definitions are used [35]:

- Full uniqueness: All profiles of all components are unique; that is, all columns of C and S are identifiable. This coincides with Definition 2 above.
- Partial uniqueness: Both profiles of one or more, but not all, components are unique.
- Fractional uniqueness: A single profile of a component is recovered uniquely, while the others are not necessarily. This coincides with Definition 3 below.
- Nonuniqueness: No profile is identifiable; that is, a unique solution does not exist even for a single profile. This definition would be different from the NMF literature, where nonuniqueness means that at least one profile is not identifiable.

In this paper, we will focus on full identifiability (Definition 2) and partial identifiability which we define as follows.

DEFINITION 3 (partial identifiability in exact NMF). *Let $R = C_\star S_\star^\top$ be an exact NMF of R of size r. The kth column of $C_\star$ is identifiable if and only if, for any other exact NMF of $R = CS^\top$ of size r, there exist an index set j and a scalar $\alpha > 0$ such that*

$$C(:,j) = \alpha C_\star(:,k).$$

Similarly, we can define the identifiability of the kth column of $S_\star$ using symmetry, which is referred to as the duality principle in the chemometrics literature [32], since $R = C_\star S_\star^\top$ if and only if $R^\top = S_\star C_\star^\top$. We will focus in this paper on the partial identifiability of the first factor, $C_\star$, without loss of generality (w.l.o.g.) by symmetry of the problem: any result that applies to $C_\star$ applies to $S_\star$.

Most results on the identifiability of exact NMF focus on the case $r = \operatorname{rank}(R)$, as it is the most reasonable in most applications. We will also focus on this case in this paper.

Contribution and outline of the paper. Although partial identifiability has been considered in the chemometrics literature, there does not exist, to the best of our knowledge, a detailed formal description (that is, a formal mathematical theorem) of the assumptions needed to obtain such results, nor do rigorous proofs exist. The main contribution of this paper is to provide several new theorems regarding the partial identifiability of exact NMF.

The paper is organized as follows: In section 2, we briefly recall the geometric interpretation of exact NMF on which our results and many identifiability results in the literature rely on. In section 3, we review important results on the identifiability

of exact NMF that will be useful in our discussions. Section 4 contains our main contributions, namely,

- the restricted data-based uniqueness (DBU) theorem (Theorem 6), a partial identifiability theorem for exact NMF;
- a geometric interpretation of the restricted DBU theorem (Lemma 1), which will lead us to a new partial identifiability theorem for exact NMF (Theorem 7);
- a new theorem allowing us to use any partial identifiability theorem sequentially to guarantee the uniqueness of several columns of C and S (Theorem 8);
- a new partial identifiability theorem for exact NMF in the special case $r = 3$ (Theorem 9).

Finally, in section 5, we discuss the practical implications of our result, provide an algorithm to automatically check partial identifiability in an exact NMF which is available at https://gitlab.com/ngillis/nmf-partial-identifiability along with all the examples presented in the paper, and illustrate the algorithm in an example from the chemometrics literature. Note that we also provide small examples throughout the paper to illustrate our theoretical results.

2. Preliminary: Geometric interpretation of exact NMF. Most results on the identifiability of exact NMF rely on its geometric interpretation, including the results of this paper. We therefore briefly recall it here for completeness.

For an exact NMF $R = CS^\top$, we can assume w.l.o.g. that R, C, and $S^\top$ are column stochastic; that is, the entries in each column sum to one. Hence each column of R, C, and $S^\top$ has unit ℓ_1-norm (a.k.a. absolute sum norm, area norm, grid norm, taxi cabnorm, Manhattan norm). The ℓ_1-norm coincides with the so-called Borgen norm with $z = e$ in the chemometrics literature, where e is the vector of all ones of appropriate dimension [16, 33]. In fact, one can first remove zero columns and rows of R and remove the corresponding columns and rows of $S^\top$ and C, respectively, which do not bring any useful information, while it may lead to numerical problems [31]. Then one can normalize $R = CS^\top$ as follows:

$$R_n(:,j) := \frac{R(:,j)}{R(:,j)^\top e} = \sum_{k=1}^{r} \underbrace{\frac{C(:,k)}{C(:,k)^\top e}}_{:=C_n(:,k)} \underbrace{\frac{C(:,k)^\top e}{R(:,j)^\top e} S(j,k)}_{:=S_n(j,k)} = \sum_{k=1}^{r} C_n(:,k) S_n(j,k). \tag{2}$$

Hence $R_n = C_n S_n^\top$, where R_n and C_n are column stochastic (that is, $e^\top = e^\top R_n$ and $e^\top C_n = e^\top$) by construction, while $S_n^\top$ is because

$$e^\top = e^\top R_n = e^\top C_n S_n^\top = e^\top S_n^\top. \tag{3}$$

Let us therefore assume, w.l.o.g., that R, C, and $S^\top$ are column stochastic. See the chemometrics analogue using Borgen norms and closure in [30, 33]. This means that, after normalization, the columns of R belong the convex hull of the columns of C that are column stochastic since, for all j,

$$R(:,j) = \sum_{k=1}^{r} C(:,k) S(j,k) = CS(j,:)^\top,$$

where $S(j,:)^\top \in \Delta = \{x \mid x \geq 0, e^\top x = 1\}$, with Δ being the probability simplex of appropriate dimension. In the case $r = \text{rank}(R)$, we must have $\text{col}(R) = \text{col}(C)$, and therefore

$$\text{conv}(R) \subseteq \text{conv}(C) \subseteq \Delta \cap \text{col}(R),$$

where $\text{conv}(R) = \{x \mid x = Ry, y \in \Delta\}$; see, e.g., [12, Chapter 2]. Hence exact NMF reduces to finding a polytope (that is, a bounded polyhedron), $\text{conv}(C)$ with r vertices (the columns of C), nested between $\text{conv}(R)$ and $\Delta \cap \text{col}(R)$. This is the so-called nested polytope problem (NPP) in computational geometry which is defined as follows.

DEFINITION 4 (NPP). *Given a full-dimensional inner polytope defined by its vertices* $\{v_1, v_2, \ldots, v_n\}$*, that is,*

$$\mathcal{P}_{inn} = \text{conv}([v_1, v_2, \ldots, v_n]) \subseteq \mathbb{R}^d,$$

a full-dimensional outer polytope defined by its facets[1]

$$\mathcal{P}_{out} = \{x \in \mathbb{R}^d \mid Fx + g \geq 0\}, \text{ where } F \in \mathbb{R}^{m \times r} \text{ and } g \in \mathbb{R}^m,$$

such that $\mathcal{P}_{inn} \subseteq \mathcal{P}_{out}$*, and an integer* $p \geq d+1$*, find a polytope,* $\mathcal{P}_{bet}$*, with* p *vertices nested* between $\mathcal{P}_{inn}$ *and* $\mathcal{P}_{out}$*, that is,* $\mathcal{P}_{inn} \subseteq \mathcal{P}_{bet} \subseteq \mathcal{P}_{out}$*.*

The polytope $\text{conv}(R)$ is typically not full-dimensional since $m > r$ in most cases. In fact, $\text{conv}(R)$ has dimension $\text{rank}(R) - 1$, and the NPP corresponding to the exact NMF of R satisfies $d = \text{rank}(R) - 1$. However, up to restricting the solution space to the affine hull of R, the set $\text{conv}(R)$ plays the role of $\mathcal{P}_{inn}$ in the NPP and $\Delta \cap \text{col}(R)$ the role of $\mathcal{P}_{out}$.

THEOREM 1 ([39]). *The exact NMF problem with* $r = \text{rank}(R)$ *is equivalent to an NPP with* $d = \text{rank}(R) - 1$ *and* $p = d + 1$ *and vice versa.*

The equivalence between exact NMF and the NPP can be used to study the identifiability of exact NMF. For example, for $\text{rank}(R) = r = 2$, the NPP is trivial since $\mathcal{P}_{inn}$ and $\mathcal{P}_{out}$ are one-dimensional polytopes, that is, segments [23]; see also, e.g., [34]. The exact NMF of R when $r = 2$ is unique if and only if $\mathcal{P}_{inn} = \mathcal{P}_{out}$ in the corresponding NPP, which leads to necessary and sufficient conditions of R; see section 3.

For $r = 3$, the NPP has dimension two and has been used extensively in the MCR literature to study the identifiability of exact NMF; see, e.g., [6, 13, 36]. They refer to the NPPs with feasible regions as Borgen–Rajkó plots; see below for an example of NPPs and also section 4. In this case, it is particularly useful to know how to reduce an instance of exact NMF to the NPP and vice versa. Let us briefly recall these reductions which we will use later in the paper.

From exact NMF to the NPP. Let R be an instance of exact NMF with $r = \text{rank}(R)$. First remove zero columns and rows of R, and normalize R to become column stochastic. Let $\mathcal{L}$ be the index set of r linearly independent columns of R so that $R = R(:, \mathcal{L})V \geq 0$ for some V. Since R and $U = R(:, \mathcal{L})$ are column stochastic,

[1] A facet of a d-dimensional polytope is a $(d-1)$-dimensional face. The face of a polytope is the intersection of that polytope with any closed half space whose boundary is disjoint from the interior of the polytope. For the polytope $\mathcal{P}_{out}$, each facet will have the form $\{x \in \mathcal{P}_{out} \mid F(i,:)x + g_i = 0\}$ for some i.

the entries in each column of V sum to one by the same argument as in (3). We define $v_j = V(1:r-1,j)$ for $j = 1,2,\ldots,n$ whose convex hull forms $\mathcal{P}_{inn}$, while

$$\mathcal{P}_{out} = \{x \in \mathbb{R}^{r-1} \mid U(:,1:r-1)x + U(:,r)(1-e^\top x) \geq 0\}.$$

This NPP instance has a solution with r vertices if and only if R admits an exact NMF of size r [39].

From NPP to exact NMF. This reduction is particularly useful to construct matrices coming from NPP problems in two dimensions. Given an NPP instance, the matrix R is constructed as follows: for all $j = 1,2,\ldots,n$

$$R(:,j) = Fv_j + g, \text{ where } R(:,j) \geq 0 \text{ since } v_j \in \mathcal{P}_{inn} \subseteq \mathcal{P}_{out}.$$

The matrix R admits an exact NMF of size r if and only if the NPP instance has a solution with r vertices [39]. Observe that each row of R corresponds to a facet of $\mathcal{P}_{out}$ and each column to a vertex of $\mathcal{P}_{inn}$, while $R(i,j)$ is the so-called slack of the jth vertex with respect to the ith facet, namely, $R(i,j) = F(i,:)v_j + g_i$.

Example 1. Let us consider the NPP where $\mathcal{P}_{out}$ is the unit square $[0,1]^2$ defined with the inequalities $Fx + g \geq 0$, where

$$F = \begin{pmatrix} 0 & 0 & 1 & -1 \\ 1 & -1 & 0 & 0 \end{pmatrix}^\top, \; g = \begin{pmatrix} 0 & 1 & 0 & 1 \end{pmatrix}^\top,$$

while $\mathcal{P}_{inn}$ is the quadrilateral with the four vertices $v_1 = (0.5,0)$, $v_2 = (0,0.5)$, $v_3 = (0.25,0.75)$, and $v_4 = (0.75,0.25)$; see Figure 2 for an illustration.

The matrix R of the corresponding exact NMF problem is given by $R(:,j) = Fv_j + g$ for all j, that is,

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

Looking at Figure 2, we observe that the unique nested triangle between $\mathcal{P}_{inn}$ and $\mathcal{P}_{out}$ has the vertices $s_1 = (0,0)$, $s_2 = (1,0)$, and $s_3 = (0,1)$, implying that R has a unique exact NMF of size 3, given by

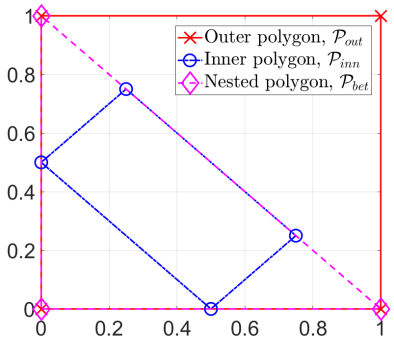

FIG. 2. *Illustration of the NPP instance described in Example* 1.

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

The first factor, C, in the above decomposition is obtained using $C(:,j) = Fs_j + g$ for $j = 1, 2, 3$.

3. Previous works on the identifiability of exact NMF. The conditions that makes exact NMF identifiable have been studied extensively in the literature. In this section, we briefly review some of the important works on the identifiability of exact NMF.

3.1. Full identifiability. Let us first discuss some conditions under which exact NMF is fully identifiable, as in Definition 2.

Necessary condition. Let us state a necessary condition for exact NMF to be fully identifiable. The condition has been rediscovered several times and is relatively easy to prove. It is based on the support of the columns of $C_\star$ and $S_\star$, the support being the set of indices containing the nonzero entries.

THEOREM 2. *Let $R = C_\star S_\star^\top$ be a fully identifiable exact NMF of R of size r. Then, the support of any column of $C_\star$ (resp., $S_\star$) does not contain the support of any other column of $C_\star$ (resp., $S_\star$).*

Proof. If a column of $C_\star$, say $C_\star(:,1)$, contains the support of another column, say $C_\star(:,2)$, then $C_\epsilon(:,1) = C_\star(:,1) - \epsilon C_\star(:,2) \geq 0$ for $\epsilon > 0$ sufficiently small, which allows us to construct another exact NMF. In fact, taking $S_\epsilon(:,2) = S_\star(:,2) + \epsilon S_\star(:,1) \geq 0$ and keeping the other columns untouched, that is, $C_\epsilon(:,k) = C_\star(:,k)$ for all $k \neq 1$ and $S_\epsilon(:,k) = S_\star(:,k)$ for all $k \neq 2$, we obtain an exact NMF $C_\epsilon S_\epsilon^\top$ which is not a permutation and scaling of $C_\star S_\star^\top$. ∎

Interestingly, this condition is also sufficient when $r = 2$, which is the only case for which we have a necessary and sufficient condition for exact NMF to be identifiable. For $r = 2$, this means that exact NMF is identifiable if and only if $C_\star$ and $S_\star$ contain a 2-by-2 diagonal submatrix: each of the two columns of $C_\star$ (resp., $S_\star$) must contain a positive entry where the other column has a zero entry.

Sufficient condition based on separability. Several identifiability results for exact NMF are based on the separability condition, defined as follows.

DEFINITION 5 (separability). *The matrix $C \in \mathbb{R}^{m \times r}$ with $m \geq r$ is separable if there exists an index set $\mathcal{K}$ of size r such that $C(\mathcal{K},:) \in \mathbb{R}^{r \times r}$ is a nonsingular diagonal matrix.*

Equivalently, the separable conditions requires that, for each $k = 1, 2, \ldots, r$, there exists an index j such that $C(j,:) = \alpha e_{(k)}^\top$ for some $\alpha > 0$, where $e_{(k)}$ is the kth unit vector (that is, the kth column of the identity matrix; recall that the notation e without the subscript $_{(k)}$ is for the all-one vector, that is, the vector of all ones of appropriate dimension). The separability condition was introduced in the NMF literature by Donoho and Stodden [10]. We have the following result.

THEOREM 3. *Let $R = C_\star S_\star^\top$ be an exact NMF of R of size r. If $C_\star$ and $S_\star^\top$ are separable, then $R = C_\star S_\star^\top$ is fully identifiable.*

It is difficult to trace back the origin of Theorem 3, and it does explicitly appear in [10], although it can be derived from their result, where they relax the condition on $C_\star$. Other sufficient conditions based on separability have been proposed in the literature, where $S_\star$ is required to be separable, while there are some sparsity conditions on $C_\star$ [21]. The results of this paper will be of this flavor but will focus on partial identifiability.

Sufficiently scattered condition. The separability condition for both factors, C and $S^\top$, is rather strong and not satisfied by most data sets; see, e.g., the discussion in [12, Chapter 4]. It can be relaxed to the following condition while retaining the full identifiability.

DEFINITION 6 (sufficiently scattered condition). *The matrix $C \in \mathbb{R}^{m\times r}$ with $m \geq r$ satisfies the sufficiently scattered condition (SSC) if there is the following:*

1. $\{x \in \mathbb{R}^r_+ \mid e^\top x \geq \sqrt{r-1}\|x\|_2\} \subseteq \text{cone}(C^\top) = \{x \mid x = C^\top h \text{ for } h \geq 0\}$.
2. *There does not exist any orthogonal matrix Q such that $\text{cone}(C^\top) \subseteq \text{cone}(Q)$, except for permutation matrices. (An orthogonal matrix Q is a square matrix such that $Q^\top Q = I$.)*

Geometrically, separability requires that $\text{cone}(C^\top)$ is the nonnegative orthant, while the SSC only requires $\text{cone}(C^\top)$ to contain the second-order (ice-cream) cone tangent to every facet of the nonnegative orthant.

THEOREM 4 ([17]). *Let $R = C_\star S_\star^\top$ be an exact NMF of R of size r. If $C_\star$ and $S_\star^\top$ satisfy the SSC, then $R = C_\star S_\star^\top$ is fully identifiable.*

It is out of the scope of this paper to discuss in detail the geometric interpretation of the SSC. An important issue with the SSC is that it is NP-hard to check in general [17]. We refer the reader to [17], [11], and [12, Chapter 4] for more details. We will briefly compare the SSC with our conditions in Remark 3.

Other full identifiability results for NMF are based on sparsity conditions; see the recent paper [1] and the references therein.

3.2. Partial identifiability. In the MCR literature, the set of feasible solutions (SFS) of exact NMF, a.k.a. the feasible regions (FRs), has been extensively studied, especially in small dimensions ($r = 3, 4$) [13]. Several algorithms exist; the best-known one is the first developed for $r = 2$ by Lawton and Sylvestre [23] who introduced and coined the special term self-modeling curve resolution (SMCR) for finding all feasible solutions for a matrix decomposition with the nonnegativity constraint. In general, the goal of MCR (and NMF) algorithms is to compute one set of particular profiles (that is, generate one solution) without considering the fact that other profiles (solutions) may exist with the same properties (namely, satisfying the same constraints and having the same objective function value). Rerunning several times these algorithms with different initializations can help to detect the nonuniqueness; however, in general, this process cannot generate the SFS. For $r = 3$, after several randomized/approximate trials, Borgen and Kowalski [6] published an analytical solution using the tangent and the simplex rotation algorithms. These algorithms were found mathematically hard to understand and implement for nonmathematicians, e.g., chemists; thus it was not developed further, although being cited in the chemometrics literature for 20 years. Rajkó and István [36] revised Borgen's study and could enlighten the concepts based on the geometry of the abstract space. Computational geometry tools (including convex hulls, Fourier–Motzkin elimination, double-description) were used for developing the algorithm to draw Borgen–Rajkó plots. The systematic grid search method

was introduced to approximate the SFS/FRs numerically first for two-component systems [40], and subsequently it was extended for three-component systems as well [14]. Sawall et al. [38] developed the polygon inflation algorithm for three-component systems as a faster and more accurate alternative to the grid search. The duality concept was first used for calculating SFS/FRs for SMCR by Beyramysoltan, Abdollahi, and Rajkó [5]. For $r = 4$, the first attempt appeared in 2013 [15] using the triangle enclosure method to approximate the boundary of the two-dimensional slices. Later in 2016, Sawall et al. [37] introduced the polyhedron inflation method as the generalization of the polygon inflation one. The chapter [37], and a subsequent paper [29] from the same research group, provided the most comprehensive summary for the SFS/FRs and related concepts up to now. See also [22] for a recent sampling algorithm for larger values of r and [2] for an improved algorithm for the boundary curve construction along with an implementation.

Necessary condition for partial identifiability. Interestingly, the necessary condition for the identifiability for exact NMF based on the supports of the columns of C can be extended to the partial identifiability case. Note that this result is, to the best of our knowledge, not present in the literature, although it follows directly from the proof of Theorem 2.

THEOREM 5. *Let $R = C_\star S_\star^\top$ be an exact NMF of R of size r. If the kth column of $C_\star$ (resp., $S_\star$) is identifiable, then the support of the kth column of $C_\star$ (resp., $S_\star$) does not contain the support of any other column of $C_\star$ (resp., $S_\star$).*

Proof. This proof is similar to that of Theorem 2. ∎

Sufficient conditions for partial identifiability: DBU and restricted DBU theorems. In the paper [35], a partial identifiability result for NMF is presented and discussed; it is called the DBU concept. Data-based means there that it does not only use the estimated profiles but also the data generated by them and all feasible profiles. It was formulated based on band solutions, that is, using not just a particular set of estimated profiles (that is, a particular exact NMF solution) but all feasible solutions based on SMCR (that is, the corresponding NPP with FRs a.k.a. Borgen–Rajkó plots [13]). Thus the SFS/FRs are needed to use the original DBU concept [35]. However, as explained in section 3.2, there are working algorithms to get SFS/FRs only for up to four-component systems. This fact inspired the development of the particular-profile DBU or restricted DBU [20] that uses a particular solution. It was a step back to the profile-based concept, which is also used in different ways by Maeder [25], Malinowski [26], and Manne [28]. The concept was intended for practitioners (such as analytical chemists) which is why both papers [20, 35] were published in *Analytica Chimica Acta*; thus the rigorous mathematical descriptions are missing. In the following, the lack of the formal descriptions and proofs will be remedied.

The idea [20] was restricted to analyze a particular solution relying on the following two conditions: Given $R = C_\star S_\star^\top$, where $C_\star \in \mathbb{R}_+^{m\times r}$ and $S_\star \in \mathbb{R}_+^{n\times r}$ with $\operatorname{rank}(R) = r$,

- (zero-region window) there exists a row of $C_\star$, say the ith, such that $C_\star(i,k) = 0$ and all feasible profiles $C(i,p) > 0$ for all $p \neq k$;
- (selective window) there exists a row of $S_\star$, say the jth, such that $S_\star(j,:) = \alpha e_{(k)}^\top$ for some $\alpha > 0$, that is, $S_\star(j,k) > 0$ and all feasible profiles $S(j,p) = 0$ for all $p \neq k$.

Let us comment on the two conditions above:

1. Zero-region window: This condition means that $C_\star(:,k)$ contains an entry equal to zero, where all other entries of $C_\star$ in the same row are positive.

Geometrically, this means that $C_\star(:,k)$ is the only column of $C_\star$ on some facet of the nonnegative orthant.

2. Selective window: This condition means that there exists a column of R, say the jth, such that $R(:,j) = \gamma C(:,k)$ for some $\gamma > 0$. In other words, it means that the kth column of C appears, up to scaling, in the data set. This is closely related to the separability condition in the NMF literature; see the previous section 3.1. In fact, all columns of $C_\star$ satisfy the selective window condition if and only if $S_\star^\top$ is separable.

As was mentioned above, the restricted DBU concept in [20] does not have a formal statement, nor a formal proof. Authors provide an informal one with explanations, focusing on the intuitions behind their result, which is more suitable for non–mathematically trained practitioners.

4. Partial identifiability theorems for exact NMF. In this section, we propose a rigorous statement and proof for the restricted DBU concept from [20]; see Theorem 6 (section 4.1). In section 4.2, we provide a geometric interpretation of Theorem 6. This leads us to a new partial identifiability result for exact NMF, Theorem 7, in section 4.3. In section 4.4, we show how to apply Theorems 6 and 7 to allow the identifiability of more than one column of $C_\star$; see Theorem 8. Finally, in section 4.5, we use the geometric interpretation of Theorem 7 to obtain a new partial identifiability result for the special case $r = 3$; see Theorem 9.

4.1. Restricted DBU theorem. Let us state and prove the restricted DBU theorem.

THEOREM 6 (restricted DBU theorem). *Let $R = C_\star S_\star^\top$, where $C_\star \in \mathbb{R}_+^{m\times r}$ and $S_\star \in \mathbb{R}_+^{n\times r}$ with $\operatorname{rank}(R) = r$. The kth column of $C_\star$ is identifiable if the following two conditions hold:*

- *(Full-rank zero-region window (FRZRW)) Let $\mathcal{I} = \{i \mid C_\star(i,k) = 0\}$ be the complement of the support of the kth column of $C_\star$. The submatrix of $C_\star$ formed by the rows indexed by $\mathcal{I}$ has rank $r-1$, that is, $\operatorname{rank}(C_\star(\mathcal{I},:)) = r-1$.*
- *(Selective window) There exists a row of $S_\star$, say the jth, such that $S_\star(j,:) = \alpha e_{(k)}^\top$ for $\alpha > 0$.*

Proof. Let $R = CS^\top$ be an exact NMF of R of size r, that is, $C \in \mathbb{R}_+^{m\times r}$ and $S \in \mathbb{R}_+^{n\times r}$. We need to show that $C(:,\ell) = \beta C_\star(:,k)$ for some ℓ and some $\beta > 0$. Since $R = CS^\top = C_\star S_\star^\top$ and since $S_\star(j,:) = \alpha e_{(k)}^\top$ for some $\alpha > 0$ (selective window condition), we have

$$R(:,j) = C_\star S_\star(j,:)^\top = \alpha C_\star(:,k) = CS(j,:)^\top = \sum_{p=1}^{r} C(:,p)S(j,p). \tag{4}$$

Let us denote the set of indices corresponding to columns of C that have zero elements in $\mathcal{I}$ as

$$\mathcal{P} = \{p \mid C(\mathcal{I},p) = 0\} \subseteq \{1,2,\ldots,r\}$$

and $\overline{\mathcal{P}} = \{1,2,\ldots,r\}\backslash\mathcal{P}$ as its complement. By nonnegativity of all the terms in (4), $S(j,p) = 0$ for all $p \in \overline{\mathcal{P}}$; otherwise a zero entry of $C_\star(:,k)$ is approximated by a positive one since $C(\mathcal{I},p) \neq 0$ for $p \in \overline{\mathcal{P}}$. Note that $|\mathcal{P}| \geq 1$; otherwise $C_\star(:,k)$ cannot be reconstructed since $C_\star(:,k) \neq 0$ as $\operatorname{rank}(C_\star) = r$. Below, we prove that $|\overline{\mathcal{P}}| \geq r-1$, and hence $|\mathcal{P}| \leq 1$. This will imply that $|\mathcal{P}| = 1$, that is, $\mathcal{P} = \{\ell\}$ for some ℓ. Putting this back into (4), this gives

$$\alpha C_\star(:,k) = \sum_{p\in\mathcal{P}=\{\ell\}} C(:,p)S(j,p) + \sum_{p\in\overline{\mathcal{P}}} C(:,p)\underbrace{S(j,p)}_{=0} = C(:,\ell)S(j,\ell),$$

where $S(j,\ell) > 0$ since $C_\star(:,k) \neq 0$ and $\alpha > 0$. Finally, $C(:,\ell) = \frac{\alpha}{S(j,\ell)} C_\star(:,k)$ which completes the proof.

It remains to show that $|\overline{\mathcal{P}}| \geq r-1$. For this, let us show that the rank of $R(\mathcal{I},:)$ is $r-1$. First, note that $\operatorname{rank}(R) = \operatorname{rank}(C_\star S_\star) = \operatorname{rank}(C_\star) = \operatorname{rank}(S_\star) = r$ by the conditions that $\operatorname{rank}(R) = r$, $R = C_\star S_\star^\top$; both $C_\star$ and $S_\star$ have r columns. Then,

$$R(\mathcal{I},:) = C_\star(\mathcal{I},:)S_\star^\top = \sum_{p\neq k} C_\star(\mathcal{I},p)S_\star(:,p)^\top = C_\star(\mathcal{I},\mathcal{K})S_\star(:,\mathcal{K})^\top,$$

where $\mathcal{K} = \{1,2,\ldots,r\}\backslash\{k\}$. By the FRZRW condition, $\operatorname{rank}(C_\star(\mathcal{I},\mathcal{K})) = r-1$, while we have $\operatorname{rank}(S_\star(:,\mathcal{K})) = r-1$ since it is made of $r-1$ columns of $S_\star$ that have rank r. Since both factors in the decomposition $R(\mathcal{I},:) = C_\star(\mathcal{I},\mathcal{K})S_\star(:,\mathcal{K})^\top$ have full rank $r-1$, $\operatorname{rank}(R(\mathcal{I},:)) = r-1$. Now, since $R = CS^\top$, we also have

$$R(\mathcal{I},:) = C(\mathcal{I},:)S^\top = C(\mathcal{I},\mathcal{P})S(:,\mathcal{P})^\top + C(\mathcal{I},\overline{\mathcal{P}})S(:,\overline{\mathcal{P}})^\top = C(\mathcal{I},\overline{\mathcal{P}})S(:,\overline{\mathcal{P}})^\top$$

since $C(\mathcal{I},\mathcal{P}) = 0$ by definition. As shown above, $\operatorname{rank}(R(\mathcal{I},:)) = r-1$. This implies that $C(\mathcal{I},\overline{\mathcal{P}})$ has at least $r-1$ columns, that is, $|\overline{\mathcal{P}}| \geq r-1$. ∎

Let us illustrate Theorem 6 in a simple example.

Example 2. Let us consider

$$R = \underbrace{\begin{pmatrix} 2 & 2 & 2 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \\ 0 & 2 & 2 \\ 0 & 1 & 2 \end{pmatrix}}_{C_\star} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{S_\star^\top}.$$

By Theorem 6, the first column of $C_\star$ is uniquely identifiable since the two conditions of Theorem 6 are satisfied.

1. (FRZRW) $C_\star(\mathcal{I},1) = 0$ for $\mathcal{I} = \{4,5\}$, while

$$\operatorname{rank}\big(C_\star(\mathcal{I},\mathcal{K})\big) = \operatorname{rank}\begin{pmatrix} 2 & 2 \\ 1 & 2 \end{pmatrix} = 2, \quad \text{where } \mathcal{K} = \{2,3\}.$$

2. (Selective window) $S(1,:) = e_{(1)}^\top$ so that $R(:,1) = C_\star(:,1)$.

Remark 1. In the example above, $S_\star$ is the identity matrix with $n = r$, which is not realistic, is not a very interesting NMF decomposition (it is the trivial decomposition, $R = RI$), and would be useless in practice. However, for our purpose, such examples are enough. One could add any number of rows to $S_\star$ and replace the identity matrix by a diagonal matrix to make it more realistic, but it would not change our observations and discussions about the identifiability.

Remark 2. The strengthened FRZRW condition compared to the zero-region window condition used in [20] comes from the fact that Theorem 6 provides a global uniqueness result. The result in [20] implicitly focuses on locally unique (a.k.a. locally

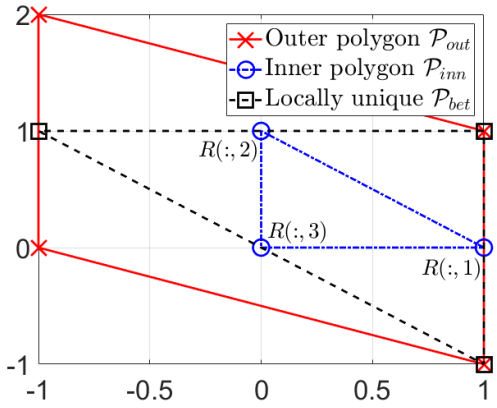

FIG. 3. *Illustration of the NPP instance corresponding to the matrix R in* (5)*. The nested polygon corresponding to the trivial factorization in* (5)*, $R = R\ I$, is $\mathcal{P}_{inn}$ itself, while the nested polygon corresponding to the factorization in* (6) *is denoted $\mathcal{P}_{bet}$.*

rigid) solutions; see [19] for more details on local uniqueness and rigidity of exact NMF solutions. For example,

$$R = \underbrace{\begin{pmatrix} 2 & 2 & 2 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \\ 0 & 2 & 2 \end{pmatrix}}_{=C_\star} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{=S_\star^\top} \tag{5}$$

satisfies the zero-region window and selective window conditions for the first column of $C_\star$: the last row, $[0,2,2]$, is a zero-region window, while the first row of $S_\star$ is ${e_{(1)}}^\top$ and hence is a selective window. However, this first column is not uniquely identifiable, up to scaling, as there exists another decomposition where that column does not appear (up to scaling):

$$R = \begin{pmatrix} 2 & 2 & 2 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \\ 0 & 2 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 \\ 1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix}. \tag{6}$$

However, in the first factorization, in (5), the first column of R is actually locally unique: any nearby exact NMF factorization must contain $R(:,1)$ as a column up to scaling. Note that, in the second factorization above, in (6), all columns of the first factor are locally partially identifiable; see Figure 3 for the corresponding NPP instance.

An interesting direction of research would be to analyze conditions under which solutions are partially locally unique.

Remark 3(FRZRW and SSC). It turns out that the FRZRW condition for each column of C is a necessary condition for the SSC. In fact, the SSC requires that C has at least $r-1$ zero per column, while the submatrix $C(\mathcal{I},\mathcal{K})$ (using the same notation as in the proof of Theorem 6) needs to contain the all-one vector in its relative interior [12, Theorem 4.28] which requires that the rank of $C(\mathcal{I},\mathcal{K})$ is equal to $r-1$.

Hence the SSC is stronger than the FRZRW condition. However, if both $C_\star$ and $S^\top$ satisfy the SSC, the exact NMF of $R = CS^\top$ is unique, which is not the case for the FRZRW condition. For example, the following matrix

$$C_\star^\top = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \end{pmatrix}$$

satisfies the FRZRW condition, but $CC^\top$ does not admit a unique NMF, e.g., $C_\star C_\star^\top = CC^\top$, where

$$C^\top = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix};$$

see [12, Example 4.29].

In the next section, section 4.2, we show that Theorem 6 has a simple geometric interpretation in terms of the NPP. This will be used to obtain a new partial identifiability result for exact NMF in section 4.3 (Theorem 7).

4.2. Geometric interpretation of the restricted DBU theorem (Theorem 6). Let us consider an NPP with $\mathcal{P}_{inn} \subseteq \mathcal{P}_{out}$, and make the following simple observation. If a vertex, v, of the inner polytope, $\mathcal{P}_{inn}$, coincides with a vertex of the outer polytope, $\mathcal{P}_{out}$, then it has to belong to any nested polytope, $\mathcal{P}_{bet}$. In fact, since $\mathcal{P}_{inn} \subseteq \mathcal{P}_{bet} \subseteq \mathcal{P}_{out}$ and $v \in \mathcal{P}_{inn} \cap \mathcal{P}_{out}$, we must have $v \in \mathcal{P}_{bet}$.

It turns out the conditions of Theorem 6 (namely, the selective window and FRZRW conditions) are equivalent to the condition that the inner and outer polytopes in the corresponding NPP share a vertex but written in algebraic terms. Let us prove this equivalence. We will then use this geometric insight to provide a new partial identifiability result in section 4.3.

In this section, we work on the outer polytope directly obtained from the reduction from exact NMF to the NPP; see section 2. It is given by

$$\mathcal{C} = \mathrm{col}(C) \cap \Delta = \{x \mid x = Cz \geq 0, e^\top x = 1\} = \{Cz \mid Cz \geq 0, e^\top z = 1\},$$

where C is normalized to be column stochastic, so that $x = Cz$ is column stochastic if and only if $e^\top z = 1$ since $e^\top x = e^\top Cz = e^\top z$. It will be useful to note that the facets of $\mathcal{C}$ have the form $\{Cz \in \mathcal{C} \mid (Cz)_i = C(i,:)z = 0\}$ for some i.

Let us define the smallest dimensional face of the outer polytope, $\mathcal{C}$, containing a given point.

Definition 7 (minimal face of $\mathcal{C}$ containing y). *Given a column stochastic matrix, $C \in \mathbb{R}_+^{m \times r}$, and the vector $y \in \mathcal{C} = \mathrm{col}(C) \cap \Delta$, we define*

$$\mathcal{F}_C(y) = \{x \in \mathcal{C} \mid \mathrm{supp}(x) \subseteq \mathrm{supp}(y)\}. \tag{7}$$

The set $\mathcal{F}_C(y)$ can be characterized as follows:

$$\mathcal{F}_C(y) = \{Cz \mid z \in \mathbb{R}^r, Cz \geq 0, (Cz)_i = 0 \text{ when } y_i = 0, z^\top e = 1\}.$$

This means that all the points in $\mathcal{F}_C(y)$ have to belong to the same facets of $\mathcal{C}$ as y. Hence $\mathcal{F}_C(y)$ is the face of $\mathcal{C}$ of minimal dimension containing y because a face of a polytope is obtained by intersecting a subset of its facets. Note that a vertex

of a poltyope is a zero-dimensional face, and hence y is a vertex of $\mathcal{C}$ if and only if $\mathcal{F}_C(y) = \{y\}$.

Let us now prove that the FRZRW condition of Theorem 6 is equivalent to the fact that $C(:,k)$ is a vertex of $\mathcal{C} = \text{col}(C) \cap \Delta$, that is, $\mathcal{F}_C(C(:,k)) = \{C(:,k)\}$. Note that the selective window assumption of Theorem 6 will require that the inner polytope has a vertex corresponding to $C(:,k)$.

LEMMA 1. *Given a nonsingular column stochastic matrix, $C \in \mathbb{R}_+^{m \times r}$, the FRZRW condition on the kth column of C is equivalent to the following geometric condition:*

$$\mathcal{F}_C\big(C(:,k)\big) \;=\; \big\{C(:,k)\big\}. \tag{8}$$

Proof. Recall that $\mathcal{I}$ denotes the set of indices corresponding to the zero entries in $C(:,k)$, and let us denote $\overline{\mathcal{I}}$ its complement which is the support of $C(:,k)$.

$\boxed{\Rightarrow}$ Assume the FRZRW condition holds. Let $x = Cz \in \mathcal{F}_C(C(:,k))$, that is, $Cz \geq 0$, $(Cz)_i = 0$ for $i \in \mathcal{I}$, $z^\top e = 1$. The condition $(Cz)_i = 0$ for $i \in \mathcal{I}$ can be written as $C(\mathcal{I},:)z = 0$. Since $C(\mathcal{I},k) = 0$, by definition, this requires $C(\mathcal{I},\mathcal{K})z(\mathcal{K}) = 0$, where $\mathcal{K} = \{1,2,\ldots,r\}\backslash\{k\}$ and the rank of $C(\mathcal{I},\mathcal{K})$ is $r-1$, by the FRZRW condition, and hence $z(\mathcal{K}) = 0$. This implies that $z = e_{(k)}$ since $z^\top e = 1$, and therefore (8) holds.

$\boxed{\Leftarrow}$ Assume (8) holds. Since C is nonsingular, (8) is equivalent to assuming that the solution to the system

$$Cz \geq 0, C(\mathcal{I},:)z = 0, e^\top z = 1,$$

is unique and given by $z = e_{(k)}$. The SFS of the above system can be written as

$$\mathcal{Z} = \{z \mid C(\overline{\mathcal{I}},:)z \geq 0, C(\mathcal{I},\mathcal{K})z(\mathcal{K}) = 0, e^\top z = 1\}.$$

Because of (8) and C being nonsingular, $\mathcal{Z} = \{e_{(k)}\}$. Since $C(\overline{\mathcal{I}},k) > 0$, by definition, $z = e_{(k)}$ belongs to the relative interior of $\mathcal{Z}$. Let us show that $\text{rank}(C(\mathcal{I},\mathcal{K})) < r-1$ implies that the relative interior of $\mathcal{Z}$ is made of more than one point, leading to a contradiction, and hence $\text{rank}(C(\mathcal{I},\mathcal{K})) = r-1$ since $|\mathcal{K}| = r-1$. Let $y \neq 0$ belong to the kernel of $C(\mathcal{I},\mathcal{K})$, that is, $C(\mathcal{I},\mathcal{K})y = 0$, with $e^\top y = \beta \in \mathbb{R}$. Let us define $z' \in \mathbb{R}^r$ as follows: $z'(\mathcal{K}) = \alpha y$ and $z'(k) = 1 - \alpha\beta$ so that $e^\top z' = 1$. For α sufficiently small, we have $C(\overline{\mathcal{I}},:)z' > 0$ since

$$C(\overline{\mathcal{I}},:)z' = \alpha C(\overline{\mathcal{I}},\mathcal{K})y + \underbrace{C(\overline{\mathcal{I}},k)}_{>0}(1-\alpha\beta),$$

and hence $z' \in \mathcal{Z}$ while $z' \neq e_{(k)}$. ∎

Lemma 1 implies that, for $r = 2$, the conditions of Theorem 6 are necessary and sufficient since the condition that $\mathcal{P}_{inn}$ and $\mathcal{P}_{out}$ have a vertex that coincides is necessary and sufficient; see section 3.

4.3. New partial identifiability theorem for exact NMF. By Theorem 2, for a column of C to be identifiable, it has to belong to at least one facet of $\mathcal{C}$ where the other columns of C are not located. In fact, its support cannot contain the support of any other column of C. Geometrically, this means that, for $C(:,k)$ to be identifiable, a necessary condition is that $\mathcal{F}_C(C(:,k))$ is a face of dimension smaller than or equal to $r-2$ (recall that $\mathcal{C}$ has dimension $r-1$), where no other column of C is located, that is,

$$C(:,j) \notin \mathcal{F}_C\big(C(:,k)\big) \quad \text{for all } j \neq k.$$

Inspired by this observation, we obtain a new sufficient condition for partial identifiability in the following theorem.

THEOREM 7. *Let* $R = C_\star S_\star^\top$, *where* $C_\star \in \mathbb{R}_+^{m\times r}$ *and* $S_\star \in \mathbb{R}_+^{r\times n}$ *with* $\operatorname{rank}(R) = r$. *W.l.o.g., assume* $R, C_\star$ *and* $S_\star^\top$ *are column stochastic; see* (3) *and* (2). *The* k*th column of* $C_\star$ *is identifiable if it satisfies the selective window condition, and there exists a subset,* $\mathcal{J}$, *of* $r-1$ *columns of* R, *namely,* $R(:,\mathcal{J})$, *such that* $\operatorname{rank}(R(:,\mathcal{J})) = r-1$ *and, for all* $j \in \mathcal{J}$,

$$\mathcal{F}_{C_\star}\big(C_\star(:,k)\big) \cap \mathcal{F}_{C_\star}\big(R(:,j)\big) \quad = \quad \emptyset; \tag{9}$$

that is, the minimal face on which the k*th columns of* $C_\star$ *lie on does not intersect the minimal faces on which the columns of* $R(:,\mathcal{J})$ *lie on.*

Proof. Let $R = CS^\top$ be another exact NMF of R of size r, where, w.l.o.g., we assume C and S are column stochastic. Let $\mathcal{K} = \{1,2,\dots,r\}\backslash\{k\}$. We have

$$R = C_\star(:,k)S_\star(:,k)^\top + C_\star(:,\mathcal{K})S_\star(:,\mathcal{K})^\top = \sum_{j=1}^{r} C(:,j)S(:,j)^\top.$$

Let us introduce the following terminology: given two nonnegative matrices, A and B, of the same dimension, we say that A touches B if there exists (i,j) such that $A(i,j) > 0$ and $B(i,j) > 0$. Below, we show that (9) implies that, for $j = 1,2,\dots,r$, it is not possible that $C(:,j)S(:,j)^\top$ touches $C_\star(:,k)S_\star(:,k)^\top$ while $C(:,j)S(\mathcal{J},j)^\top$ touches $R(:,\mathcal{J})$. Since $R(:,\mathcal{J})$ has rank $r-1$, by the exclusion principle, exactly one rank-one factor touches $C_\star(:,k)S_\star(:,k)^\top$, and hence it has to coincide with it.

Assume $C(:,p)S(:,p)^\top$ touches $C_\star(:,k)S_\star(:,k)^\top$ and $C(:,p)S(\mathcal{J},p)^\top$ touches $R(:,\mathcal{J})$ for some p. As $\operatorname{rank}(R(:,\mathcal{J})) = r-1$, $R(:,j) \neq 0$ for all $j \in \mathcal{J}$. Since $C(:,p)S(:,p)^\top$ touches $C_\star(:,k)S_\star(:,k)^\top$, the support of $C(:,p)$ is contained in the support of $C_\star(:,k)$, and hence $C(:,p) \in \mathcal{F}_{C_\star}(C_\star(:,k))$. By (9), $C(:,p) \notin \mathcal{F}_{C_\star}(R(:,j))$ for all $j \in \mathcal{J}$; that is, the support of $C(:,p)$ is not contained in the support of any column of $R(:,\mathcal{J})$ implying that it cannot touch any column of $R(:,\mathcal{J})$, a contradiction. ∎

The condition in Theorem 7 implies that the kth column of $S_\star$ contains at least $r-1$ entries equal to zero, namely, $S_\star(\mathcal{J},k) = 0$, since (9) requires that the support of $R(:,j)$ for $j \in \mathcal{J}$ does not contain the support of $C_\star(:,k)$.

Example 3. Let us consider the NPP where $\mathcal{P}_{out}$ is the unit square $[0,1]^2$ as in Example 1, while $\mathcal{P}_{inn}$ is the triangle with the vertices $v_1 = (0.5,0)$, $v_2 = (0.2,1)$, and $v_3 = (0.8,1)$; see Figure 4 for an illustration.

The matrix R of the corresponding exact NMF problem is given by $R(:,j) = Fv_j + g$ for all j, that is,

$$R = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0.5 & 0.2 & 0.8 \\ 0.5 & 0.8 & 0.2 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0.5 & 0 & 1 \\ 0.5 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.8 & 0.2 \\ 0 & 0.2 & 0.8 \end{pmatrix}.$$

The first column of C satisfies the selective window assumption, while we observe in Figure 4 that Theorem 7 applies using $R(:,[2,3])$ whose minimal faces do not intersect with that of $C_\star(:,1)$ which is therefore identifiable. Note that the restricted DBU Theorem 6 is not applicable to $C_\star(:,1)$ since it does not correspond to a vertex of $\mathcal{P}_{out}$.

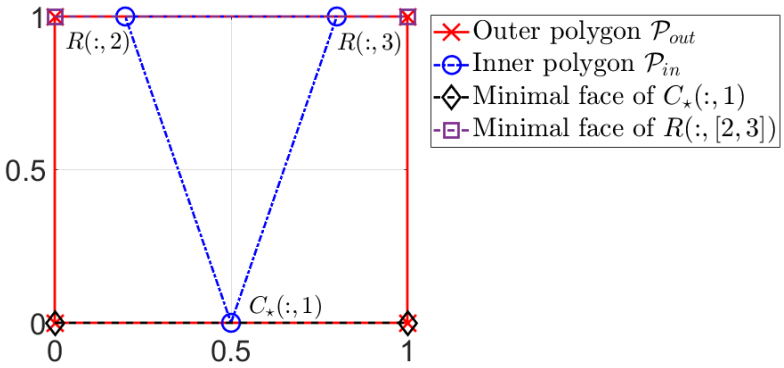

FIG. 4. *Illustration of the NPP instance described in Example* 3.

It is important to note the following:

- Theorem 7 does not subsume Theorem 6 which applies to a column of C which is a vertex of $\mathcal{P}_{out}$ in which case the existence of a subset of columns of R satisfying (9) is not necessary. For example, taking $\mathcal{P}_{out}$ as the square in two dimensions, as above, and taking the vertices of $\mathcal{P}_{inn}$ as $v_1 = (0,0)$ (bottom left corner), $v_2 = (0, 0.5)$, and $v_3 = (0.5, 0)$, the conditions of Theorem 7 do not apply to the first column of C (corresponding to v_1 since the minimal faces of v_2 and of v_3 contain v_1), while Theorem 6 does apply.
- For condition (9) to be satisfied, a necessary, but not sufficient, condition is that the supports of $C_\star(:,k)$ and $R(:,j)$ are not contained in one another. In fact, this support condition implies that $\mathcal{F}_{C_\star}(C_\star(:,k))$ and $\mathcal{F}_{C_\star}(R(:,j))$ are distinct faces but not that their intersection is empty.

Theorem 7 can be directly used to obtain a full identifiability result.

COROLLARY 1. *Let* $R = C_\star S_\star^\top$, *where* $C_\star \in \mathbb{R}_+^{m\times r}$ *and* $S_\star \in \mathbb{R}_+^{r\times n}$ *with* rank(R) = r. *W.l.o.g., assume* $R, C_\star$ *and* $S_\star^\top$ *are column stochastic. If*

1. *every column of* $C_\star$ *satisfies the selective window assumption, that is,* $S_\star^\top$ *is separable, and*
2. *the following holds for all* $k \neq j$,

$$\mathcal{F}_{C_\star}\big(C_\star(:,k)\big) \cap \mathcal{F}_{C_\star}\big(C_\star(:,j)\big) \quad = \quad \emptyset,$$

then $(C_\star, S_\star)$ *is (fully) identifiable.*

Proof. On one hand, Theorem 7 applies to all columns of $C_\star$, taking $R(:,\mathcal{J}) = C(:,\{1\ldots,r\}\backslash\{k\})$ for all $k = 1,2,\ldots,r$, since $S_\star^\top$ is separable. On the other hand $S_\star$ is identifiable since $C_\star$ is identifiable and $\text{rank}(C_\star) = r$. ∎

For example, in two dimensions, when $r = 3$, full identifiability based on Corollary 1 requires that, in the NPP, the three vertices of $\mathcal{P}_{inn}$ corresponding to the three columns of $C_\star$ are located on three nonadjacent edges of the polygon $\mathcal{P}_{out}$. Note that this requires $\mathcal{P}_{out}$ to have at least six edges, that is, to be an n-gon with $n \geq 6$. This implies that R needs to have at least 6 rows.

Example 4. Let us take an example with $r = 4$, for which the NPP has dimension three. Consider the outer polytope as the unit cube in dimension three, with $\mathcal{P}_{out} =$

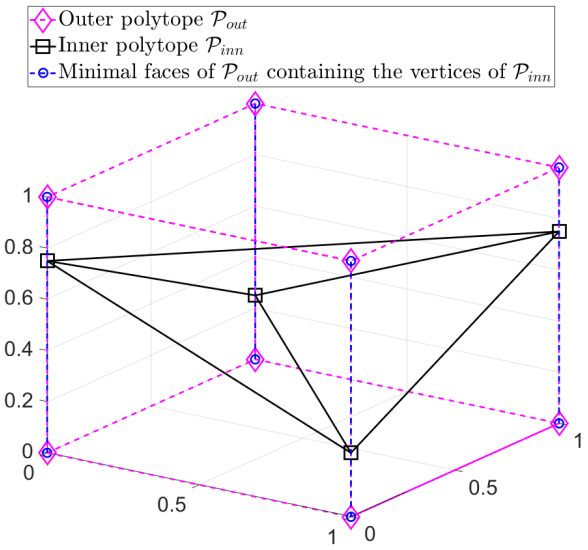

FIG. 5. *Geometric interpretation of Example* 4 *that satisfies the conditions of Theorem* 7.

$[0,1]^3$, and take the vertices of $\mathcal{P}_{inn}$ as (0,0,0.75), (1, 0, 0.25), (0, 1, 0.25), and (1, 1, 0.75). This construction satisfies the conditions of Theorem 7 for all k since the vertices of $\mathcal{P}_{inn}$ are on (minimal) faces (namely, edges) that do not intersect; see Figure 5 for an illustration.

The corresponding R is given by

	$(0,0,0.75)$	$(1,0,0.25)$	$(0,1,0.25)$	$(1,1,0.75)$
$x_1 \geq 0$	0	1	0	1
$x_2 \geq 0$	0	0	1	1
$x_3 \geq 0$	0.75	0.25	0.25	0.75
$x_1 \leq 1$	1	0	1	0
$x_2 \leq 1$	1	1	0	0
$x_3 \leq 1$	0.25	0.75	0.75	0.25

and therefore has a unique exact NMF, $R = RI$. Note that no column of R satisfies the FRZRW condition of Theorem 6 since R has only two zero entries per column: Geometrically, no vertices of $\mathcal{P}_{inn}$ are a vertex of $\mathcal{P}_{out}$.

4.3.1. Is it easy to check the conditions of Theorem 7?. Let $R = CS^\top$ be an exact NMF of size $r = \operatorname{rank}(R)$, where R and C are column stochastic (w.l.o.g.). A column of R, say the jth, fails to satisfy condition (9) if and only if there exists x such that

$$x \quad \in \quad \mathcal{F}_{C_\star}\big(C_\star(:,k)\big) \cap \mathcal{F}_{C_\star}\big(R(:,j)\big).$$

Such an x exists if the following linear system in variable $z \in \mathbb{R}^r$ has a solution:

$$x = Cz \geq 0,\ z^\top e = 1,\ (Cz)_i = 0 \text{ for all } i \in \mathcal{K}_{k,j} = \{p \mid C(p,k) = 0 \text{ or } R(p,j) = 0\}.$$

This is a linear system in r variables, with $\mathcal{O}(m)$ equalities and inequalities. In our implementation (see section 5.2), to avoid numerical issue, we rather solve the linear optimization problem (which is always feasible)

$$\min_{z} \sum_{i \in \mathcal{K}_{k,j}} (Cz)_i \quad \text{such that} \quad Cz \geq 0 \text{ and } e^\top z = 1 \tag{10}$$

and check whether the optimal objective function value is below a given threshold (we used 10^{-6}).

4.4. Using partial identifiability theorems sequentially. In this section, we provide a simple general framework to generalize partial identifiability theorems, assuming a subset of columns of $C_\star$ is already identifiable.

THEOREM 8. *Let $R = C_\star {S_\star}^\top$, where $C_\star \in {\mathbb{R}^{m \times r}}_+$ and $S_\star \in {\mathbb{R}^{r \times n}}_+$ with $\operatorname{rank}(R) = r$. Assume p columns of $C_\star$ are identifiable for $p \in \{1, 2, \ldots, r-1\}$, say the first p w.l.o.g.; that is, $C_\star(:, j)$ are identifiable for $j = 1, 2, \ldots, p$ (Definition 3). Let $\mathcal{J}$ be the index set corresponding to the columns of R that do not contain the support of the first p columns of $C_\star$.*

If $\operatorname{rank}(S_\star(\mathcal{J}, p+1:r)) = r - p$ and if the $(p+1)$th column of $C_\star$ can be certified to be identifiable in the exact NMF $R(:, \mathcal{J}) = C_\star(:, p+1:r) S_\star(\mathcal{J}, p+1:r)^\top$ of size $r - p$, then $C_\star(:, p+1)$ is identifiable in the exact NMF of R of size r.

Proof. Let $R = CS^\top$ be an exact NMF of X of size r with $C \in \mathbb{R}^{m \times r}_+$ and $S \in \mathbb{R}^{n \times r}_+$. W.l.o.g., $C(:, 1:p) = C_\star(:, 1:p) D$, where D is a diagonal matrix since the first p columns of $C_\star$ are identifiable. We have

$$R(:, \mathcal{J}) = C_\star S_\star(\mathcal{J}, :)^\top = \sum_{q=p+1}^{r} C(:, q) S(\mathcal{J}, q)^\top. \tag{11}$$

The last equality follows by construction: the columns of $R(:, \mathcal{J})$ do not contain the support of the columns of $C_\star(:, 1:p)$, which coincide with that of $C(:, 1:p)$, implying $S(\mathcal{J}, q) = 0$ for all $q \leq p$. The fact that $\operatorname{rank}(S_\star(p+1:r, \mathcal{J})) = r - p$ implies that $\operatorname{rank}(R(:, \mathcal{J})) = r - p$ since $\operatorname{rank}(C(:p+1:r)) = r - p$ as $\operatorname{rank}(C) = r$, and hence (11) is an exact NMF of rank $r - p$. By assumption, $C_\star(:, p+1)$ is identifiable in the exact NMF (11) so that one of the columns of $C(:, p+1:r)$ is equal to $C_\star(:, p+1)$ up to scaling. ∎

Let us illustrate Theorem 8 in a simple example where all columns of $C_\star$ can be certified to be identifiable, using Theorem 6 sequentially.

Example 5. Let

$$R = \underbrace{\begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 2 & 1 \\ 1 & 0 & 1 & 2 \\ 1 & 0 & 2 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}}_{C_\star} \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{S_\star^\top}.$$

All columns of $C_\star$ are identifiable. The first one is by Theorem 6. The second one is by combining Theorem 8 and Theorem 6: the last three columns of R do not belong to the support of $C_\star(:, 1)$; we have

$$R(:,2:4) = \underbrace{\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 2 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}}_{C_\star(:,2:4)} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{S_\star(2:4,2:4)^\top},$$

where $\operatorname{rank}(S_\star(2:4,2:4)) = 3$. We can therefore apply Theorem 6 to the above exact NMF of size $r-p=3$, which certifies the identifiability of $C_\star(:,2)$ (the selective window and FRZRW conditions hold). One can certify the identifiability of the last two columns of $C_\star$ in the same way.

It is important to note that the conditions of Theorem 8 do not necessarily become milder as p increases. In practice, this means one needs to check $\sum_{p'=0}^{p} \binom{r}{p'}$ cases for each column of $C_\star$ not identified yet. However, this can be implemented relatively easily using recursion; see section 5.2 for the details. Let us illustrate this in another example.

Example 6. Let

$$R = \underbrace{\begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 2 & 1 \\ 1 & 0 & 1 & 2 \\ 1 & 0 & 2 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}}_{C_\star} \underbrace{\begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{S_\star^\top} = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 2 & 1 \\ 1 & 0 & 2 & 3 \\ 1 & 0 & 3 & 2 \\ 1 & 0 & 1 & 2 \\ 1 & 1 & 2 & 1 \end{pmatrix}.$$

As in Example 5, the first column of $C_\star$ is identifiable. Now, we realize that Theorem 8 with $p=1$ for the second column is not applicable: the last two columns of R do contain the support of $C_\star(:,1)$ so that $\mathcal{J}=\{2\}$, and $\operatorname{rank}(S(p+1:r,\mathcal{J}) = 1 < r-p-1 = 2$. However, the second column of $C_\star$ satisfies the conditions of Theorem 6 and hence is identifiable.

Note that the last two columns of $C_\star$ do not satisfy the selective window assumption, and it turns out that they are not identifiable since another exact NMF is given by $R = RI$.

4.5. Partial identifiability for exact NMF when $r=3$. We now analyze the case when $r=3$, which is of particular interest in the MCR literature, by providing a new condition for identifiability of two columns of $C_\star$. Before that, let us show the following lemma.

LEMMA 2. *Let $R = C_\star S_\star^\top$ be an exact NMF of R of size $r = \operatorname{rank}(R)$, where the kth column of $C_\star$ satisfies the selective window assumption. Let $R = CS^\top$ be an exact NMF of R of size r. W.l.o.g., assume $C_\star$ and C are column stochastic. If the kth column of $C_\star$ is not identified in C, that is, $C(:,j) \neq C_\star(:,k)$ for all j, then there exists an index set $\mathcal{J}$ with $|\mathcal{J}| \geq 2$ such that*

$$C(:,j) \in \mathcal{F}_{C_\star}\big(C_\star(:,k)\big) \text{ for } j \in \mathcal{J}.$$

Proof. Since $C_\star(:,k)$ satisfies the selective window assumption, that is, $C_\star(:,k) = \alpha R(:,j)$ for some j and $\alpha > 0$, we have $C_\star(:,k) = Cz$ for some $z \in \Delta$. The result then follows from the two observations:

- Since $C_\star(:,k) = Cz$, $\text{supp}(C(:,j)) \subseteq \text{supp}(C_\star(:,k))$ for all j such that for $z_j > 0$. Therefore $C(:,j) \in \mathcal{F}_{C_\star}(C_\star(:,k))$ since $\text{col}(C) = \text{col}(C_\star) = \text{col}(R)$.
- Let $\mathcal{J} = \{j \mid z_j > 0\}$. If $|\mathcal{J}| = 1$, $C(:j) = C_\star(:,k)$ for some j, a contradiction; hence $|\mathcal{J}| \geq 2$.

∎

THEOREM 9. *Let $R = C_\star S_\star^\top$, where $C_\star \in \mathbb{R}_+^{m\times 3}$ and $S_\star \in \mathbb{R}_+^{3\times n}$ with $\text{rank}(R) = 3$, and R, $C_\star$, and $S_\star^\top$ are normalized to be column stochastic as in (2). Let us assume that two columns of $C_\star$ satisfy the selective window assumption, say the first and second one w.l.o.g.. Also let the supports of $C_\star(:,1)$ and $C_\star(:,2)$ not be contained in one another. Then, these two columns are identifiable if there exists a column of R, say the jth, such that*

if $\mathcal{F}_{C_\star}(C_\star(:,2)) \cap \mathcal{F}_{C_\star}(C_\star(:,1)) = \emptyset$,

$$R(:,j) \quad \notin \quad \text{conv}\Big(\big[C_\star(:,1), \mathcal{F}_{C_\star}\big(C_\star(:,2)\big)\big]\Big) \cup \text{conv}\Big(\big[C_\star(:,2), \mathcal{F}_{C_\star}\big(C_\star(:,1)\big)\big]\Big), \tag{12}$$

else

$$R(:,j) \quad \notin \quad \text{conv}\Big(\mathcal{F}_{C_\star}\big(C_\star(:,1)\big), \mathcal{F}_{C_\star}\big(C_\star(:,2)\big)\Big). \tag{13}$$

Proof. Let $R = CS^\top$ be an exact NMF of R of size $r = 3$. The proof mostly relies on Lemma 2: if $C_\star(:,k)$ is not identified, then there are least two columns of C in $\mathcal{F}_{C_\star}(C_\star(:,k))$. Note that, for $r = 3$, $\mathcal{C}$ is a polygon, and hence there are three types of facets depending on their dimension: the two-dimensional polygon itself, $\mathcal{C}$, one-dimensional segments, and zero-dimensional vertices. By (12) or (13), $\mathcal{F}_{C_\star}(C_\star(:,j))$ for $j = 1,2$ cannot be the polygon itself and hence is either segments or vertices.

Case 1: $\mathcal{F}_{C_\star}(C_\star(:,2)) \cap \mathcal{F}_{C_\star}(C_\star(:,1)) = \emptyset$. Since C has three columns, there cannot be four columns of C in $\mathcal{F}_{C_\star}(C_\star(:,k))$ for $k \in \{1,2\}$, and therefore $C_\star(:,k)$ is identified for $k = 1$ or $k = 2$, say $C_\star(:,1)$ w.l.o.g.. Then, because of (12), $C_\star(:,2)$ must also be identified; otherwise $R(:,j)$ cannot be reconstructed. In fact, if $C_\star(:,2)$ was not identified, the two columns of C not multiples of $C_\star(:,1)$ (which is identified) must be on $\mathcal{F}_{C_\star}(C_\star(:,2))$, a contradiction between the fact that $R(:,j) = CS(j,:)^\top$ and (12).

Case 2: $\mathcal{F}_{C_\star}(C_\star(:,2)) \cap \mathcal{F}_{C_\star}(C_\star(:,1)) \neq \emptyset$. The two facets $\mathcal{F}_{C_\star}(C_\star(:,1))$ and $\mathcal{F}_{C_\star}(C_\star(:,2))$ intersect in a vertex. In fact, for $r = 3$, $\mathcal{F}_{C_\star}(C_\star(:,1))$ and $\mathcal{F}_{C_\star}(C_\star(:,2))$ are adjacent segments of $\mathcal{C}$ since the support of $C_\star(:,1)$ does not contain and is not contained in that of $C_\star(:,2)$. Moreover, by the same support condition, $C_\star(:,1) \notin \mathcal{F}_{C_\star}(C_\star(:,2))$ and vice versa. Therefore, if $C_\star(:,1)$ or $C_\star(:,2)$ is not identified, the three columns of C belong to $\mathcal{F}_{C_\star}(C_\star(:,1)) \cup \mathcal{F}_{C_\star}(C_\star(:,2))$, which is a contradiction since $R(:,j)$ does not belong to the convex hull of these sets (see (13)), and hence C cannot be used to reconstruct $R(:,j)$. ∎

Example 7. Let us construct two examples to illustrate the two cases in Theorem 9. To do so, we use the equivalence of exact NMF with the NPP and use the same outer polygon $\mathcal{P}_{out} = [0,1]^2$ as in Example 1.

In the first case of Theorem 9, the two minimal faces containing $C_\star(:,1)$ and $C_\star(:,2)$ do not intersect. For example, one can take the two points (0.5,0) and (0.5,1); see Figure 6.

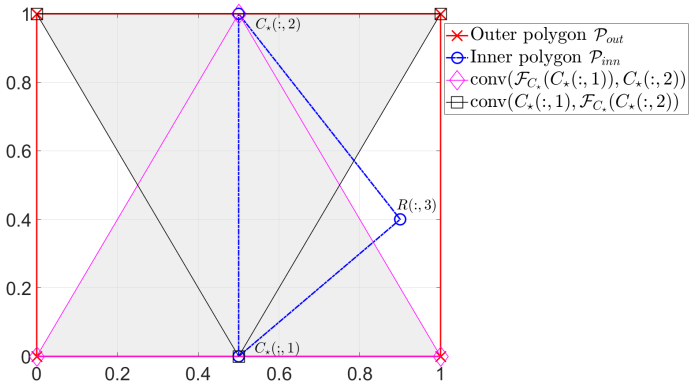

FIG. 6. *Geometric interpretation of the exact NMF problem: identifiability of the first two columns of $C_\star$, case* 1 *of Theorem* 9.

These two points correspond to

$$C_\star(:,1) = F(0.5,0) + g = (0,1,0.5,0.5)^\top \text{ and } C_\star(:,2) = F(0.5,1) + g = (1,0,0.5,0.5)^\top.$$

If a column of $R = C_\star S_\star^\top$ does not belong to

$$\text{conv}\Big(C_\star(:,1), \mathcal{F}_{C_\star}\big(C_\star(:,2)\big)\Big) \cup \text{conv}\Big(C_\star(:,2), \mathcal{F}_{C_\star}\big(C_\star(:,1)\big)\Big),$$

then both columns are identifiable. This is the case in Figure 6 with

$$R(:,3) = F(0.9,0.4) + g = (0.4,0.6,0.9,0.1)^\top.$$

In the second case of Theorem 9, the two minimal faces containing $C_\star(:,1)$ and $C_\star(:,2)$ do intersect. For example, one can take the two points (0.5,0) and (0,0.5); see Figure 7. These two points correspond to

$$C_\star(:,1) = F(0.5,0) + g = (0,1,0.5,0.5)^\top \text{ and } C_\star(:,2) = F(0,0.5) + g = (0.5,0.5,0,1)^\top.$$

If a column of $R = C_\star S_\star^\top$ does not belong to

$$\text{conv}\Big(\mathcal{F}_{C_\star}\big(C_\star(:,1)\big), \mathcal{F}_{C_\star}\big(C_\star(:,2)\big)\Big),$$

then both columns are identifiable.

This is the case in Figure 7, with

$$R(:,3) = F(0.75,0.75) + g = (0.75,0.25,0.75,0.25)^\top.$$

5. Applications of the new partial identifiability results. In this section, we first discuss whether the conditions of our identifiability results are reasonable in practice. Then we propose an algorithm, Algorithm 1, that combines our partial identifiability results to certify the partial identifiability results for a given input matrix R. Finally, we illustrate its use in an example from the chemometrics literature.

5.1. Are the conditions of our identifiability results reasonable?. All our proposed identifiability results rely on two facts:

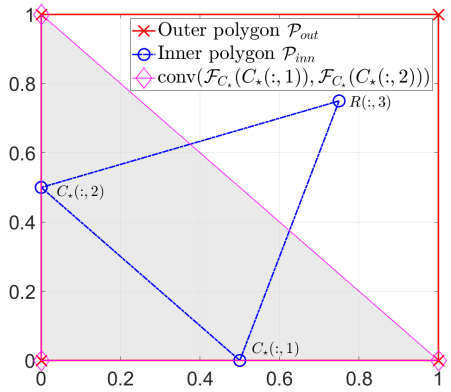

FIG. 7. *Geometric interpretation of the exact NMF problem: identifiability of the first two columns of $C_\star$, case 2 of Theorem 9.*

1. Some columns of C satisfy the selective window assumption; this requires some rows of S to be unit vectors (up to scaling).
2. These columns of C must have some degree of sparsity. (Note that this is a necessary condition for identifiability of exact NMF; see Theorem 5).

This implies some degree of sparsity in $R = CS^\top$ since some columns of R will be equal to the columns of C that have zero entries.

The selective window assumption is reasonable in many applications; see, e.g., the discussion in [12, Chapter 7] about separability and the references therein. However, sparsity is not necessarily natural in all applications where separability arises, e.g., in blind hyperspectral unmixing, where spectral signatures are typically dense, and in facial feature extraction, where facial images are dense. However, it is reasonable in other applications. The following are examples:

- MCR: The spectral content of some sources/components can be high (overlapping), while it is zero/small for others at some wavelength (selective window assumption). Moreover, all components are not present at all time window (sparsity); see an example in section 5.3.
- Topic modeling: The presence of anchor words, which are words associated to a single topic, is a reasonable assumption [4] (selective window), while most documents only discuss a few topics (sparsity). For example, using the widely used data set tdt2˙top30 (9394 documents and 19528 words) we computed an approximate exact NMF of the form $R \approx \tilde{R} = CS^\top$ for $r \in \{1, 2, \ldots, 100\}$ using a separable NMF algorithm, namely, the successive projection algorithm [3], one of the most widely used ones. All decompositions $\tilde{R} = CS^\top$ obtained are certified to be unique using the restricted DBU theorem (Theorem 6). (Here we can only certify that the exact NMF of the approximation is identifiable since there does not exist an exact NMF of R for a small r; in fact,[2] $\text{rank}(R) \geq 800$. This is often the case in practice because of the noise and model misfit.)

[2]We stopped the modified Gram–Schmidt with column pivoting at $r = 800$ after about 5 hours on a standard laptop.

Algorithm 1 Partial identifiability guarantees for C in an exact NMF $R = CS^\top$ of size rank(R).

Input: An exact NMF of $R{=}CS^\top$, with $C{\in}\mathbb{R}^{m\times r}_+$ and $S{\in}\mathbb{R}^{n\times r}_+$, where $r{=}\text{rank}(R)$.
Output: A subset $\mathcal{K}$ of the columns of C that are guaranteed to be identifiable (Definition 3).

Normalize $(R, C, S^\top)$ so that they are column stochastic; see (2).
Initialize $\mathcal{K} = \emptyset$.
Let $\mathcal{L}$ be the set of columns of C that satisfies the selective window assumption, that is, $\mathcal{L} = \{i \mid \text{there exist } k \text{ and } \alpha > 0 \text{ such that } S(k,:) = \alpha e^\top_{(i)}\}$.
% Use Theorems 6 *and* 7
for every index in $k \in \mathcal{L}\backslash\mathcal{K}$ **do**
  if $\text{rank}(C(\mathcal{I},:)) = r-1$, where $\mathcal{I} = \{i \mid C(i,k) = 0\}$ **then**
    $\mathcal{K} \leftarrow \mathcal{K} \cup \{k\}$.
  end if
  if there exists $\mathcal{J}$ s.t. $\mathcal{F}_C(C(:,k)) \cap \mathcal{F}_C(R(:,j)) = \emptyset$ for all $j \in \mathcal{J}$, $\text{rank}(R(:,\mathcal{J})) = r-1$ **then**
    $\mathcal{K} \leftarrow \mathcal{K} \cup \{k\}$.
  end if
end for
% Use Theorem 8 *combined Theorems* 6 *and* 7*, recursively*
$i = 1$
while $i \leq |\mathcal{K}|$ **do**
  $\mathcal{P} = \{1, 2, \ldots, r\}\backslash\{\mathcal{K}(i)\}$ *%* $\mathcal{K}(i)$ *is the* i*th element in the set* $\mathcal{K}$
  Let $\mathcal{J}$ be the subset of columns of R not containing the support of $C(:, \mathcal{K}(i))$.
  if $\text{rank}(S(\mathcal{J},\mathcal{P})) = r-1$ **then**
    $\mathcal{K}' =$ Algorithm 1$(C(:,\mathcal{P}), S(\mathcal{J},\mathcal{P}))$
    $\mathcal{K} \leftarrow \mathcal{K} \cup \mathcal{P}(\mathcal{K}')$,
  end if
  $i \leftarrow i + 1$
end while
if $r = 3$ **then** use Theorem 9 for pairs of indices in $\mathcal{K}$.

In summary, our results will likely apply when R contains some columns with sufficiently many zero entries, while the selective window assumption makes sense.

5.2. An algorithm to check partial identifiability. Relying on our new theoretical results, we provide in this section an algorithm that provides partial identifiability guarantees for the exact NMF of a given nonnegative matrix R; see Algorithm 1. As for all the results of this paper, Algorithm 1 assumes $\text{rank}(R) = \text{rank}(C) = r$ which is reasonable in most real-world applications. Algorithm 1 is available at https://gitlab.com/ngillis/nmf-partial-identifiability along with all the examples presented in the paper (and two other ones).

Remark 4(use of Algorithm 1 for real-world data). NMF algorithms may return C and S with many entries close to zero but not exactly zero (e.g., if the algorithm has not converged). Therefore, to check whether your computed solution is close to being (partially) identifiable, you can set these entries to zero using some threshold strategy and then call Algorithm 1.

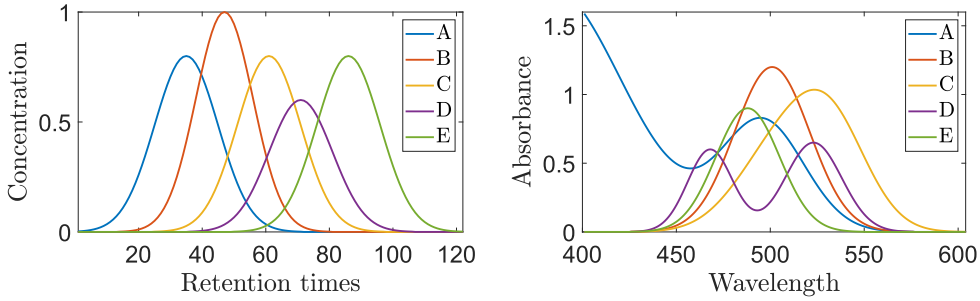

FIG. 8. *A five-component data set. On the left, the elution profiles of the chemical components which are the columns of $C^\star$. On the right, the spectra of the chemical components which are the columns of $S^\star$.*

Another strategy is to weaken the sharp zero condition in the sense of generalized Borgen plots [18]. This strategy is useful for experimental noisy data which may also include a background subtraction resulting in small negative entries.

5.3. Numerical example from the chemometrics literature. Let us consider the five-component data set from [20, section 4.3]; see Figure 8.

Algorithm 1 certifies that the first and third columns of $C^\star$ are identifiable and the fifth column of $S^\star$

`[K,L] = check_partial_identif(C,S),`
`K = [1 3], L=5.`

For example, for $C^\star$, using the restricted DBU theorem, the elution profiles (that is, the columns of $C^\star$) that can be guaranteed to be identifiable are the first (A) and third ones (C): they satisfy the selective window condition (first wavelengths for A, last ones for C), while the FRZRW condition can be checked (the other elution profiles have rank $r-1$ when restricted to the entries where the corresponding column of $C^\star$ is zero). Note that the first columns of $S^\star$ satisfy the selective window condition but not the FRZRW condition because, when its spectrum is equal to zero, the spectrum of other components also are (namely, all of them but C). These are the same conclusions as in [20].

6. Conclusion. In this paper, we have provided the following partial identifiability results for exact NMF:

- a rigorous description and proof of the restricted DBU theorem (Theorem 6);
- a new partial identifiability result based on the geometric interpretation of the restricted DBU theorem (Theorem 7);
- a sequential approach to guarantee the identifiability of more factors (Theorem 8).

Since this paper is, to the best of our knowledge, the first to rigorously investigate partial identifiability of exact NMF, there is still a lot to be done. In particular, can stronger partial identifiability theorems be obtained? For example, is it possible to provide partial identifiability results for several components simultaneously under weaker conditions? We have done this for the case $r=3$ considering two components at a time (see Theorem 9), and this idea can probably be generalized to larger r. In particular, considering all factors allows one to relax the selective window assumption (a.k.a. separability, which is rather strong) to the SSC; see Theorem 4.

Acknowledgments. We thank the reviewers for their insightful comments that helped us improve the paper.

REFERENCES

[1] M. Abdolali and N. Gillis, *Simplex-structured matrix factorization: Sparsity-based identifiability and provably correct algorithms*, SIAM J. Math. Data Sci., 3 (2021), pp. 593–623.

[2] T. Andersons, M. Sawall, and K. Neymeyr, *Analytical enclosure of the set of solutions of the three-species multivariate curve resolution problem*, J. Math. Chem., 60 (2022), pp. 1750–1780.

[3] U. Araújo, B. Saldanha, R. Galvão, T. Yoneyama, H. Chame, and V. Visani, *The successive projections algorithm for variable selection in spectroscopic multicomponent analysis*, Chemometr. Intell. Lab. Syst., 57 (2001), pp. 65–73.

[4] S. Arora, R. Ge, Y. Halpern, D. Mimno, A. Moitra, D. Sontag, Y. Wu, and M. Zhu, *A practical algorithm for topic modeling with provable guarantees*, in Proceedings of the 30th International Conference on Machine Learning, 2013, pp. 280–288.

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

[6] O. S. Borgen and B. R. Kowalski, *An extension of the multivariate component-resolution method to three components*, Analy. Chim. Acta, 174 (1985), pp. 1–26.

[7] S. Brown, R. Tauler, and B. Walczak, eds., *Comprehensive Chemometrics Chemical and Biochemical Data Analysis*, 2nd ed., Elsevier, New York, 2020.

[8] A. Cichocki, R. Zdunek, A. H. Phan, and S.-i Amari, *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-Way Data Analysis and Blind Source Separation*, John Wiley and Sons, Hoboken, NJ, 2009.

[9] P. Common and C. Jutten, *Handbook of Blind Source Separation: Independent Component Analysis and Applications*, Academic Press, Cambridge, MA, 2010.

[10] D. Donoho and V. Stodden, *When does non-negative matrix factorization give a correct decomposition into parts?*, in Advances in Neural Information Processing Systems 16 (NIPS 2003), MIT Press, Cambridge, MA, 2004, pp. 1141–1148.

[11] X. Fu, K. Huang, N. D. Sidiropoulos, and W.-K. Ma, *Nonnegative matrix factorization for signal and data analytics: Identifiability, algorithms, and applications*, IEEE Signal Process. Mag., 36 (2019), pp. 59–80.

[12] N. Gillis, *Nonnegative Matrix Factorization*, SIAM, Philadelphia, 2020.

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

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

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

[16] B.-V. Grande and R. Manne, *Use of convexity for finding pure variables in two-way data from mixtures*, Chemometr. Intell. Lab. Syst., 50 (2000), pp. 19–33.

[17] K. Huang, N. D. Sidiropoulos, and A. Swami, *Non-negative matrix factorization revisited: Uniqueness and algorithm for symmetric decomposition*, IEEE Trans. Signal Process., 62 (2013), pp. 211–224.

[18] A. Jürss, M. Sawall, and K. Neymeyr, *On generalized Borgen plots.* I*: From convex to affine combinations and applications to spectral data*, J. Chemometr., 29 (2015), pp. 420–433.

[19] R. Krone and K. Kubjas, *Uniqueness of nonnegative matrix factorizations by rigidity theory*, SIAM J. Matrix Anal. Appl., 42 (2021), pp. 134–164.

[20] M. A. Lakeh, H. Abdollahi, and R. Rajkó, *Predicting the uniqueness of single non-negative profiles estimated by multivariate curve resolution methods*, Analyt. Chim. Acta, 1199 (2022), 339575.

[21] H. Laurberg, M. G. Christensen, M. D. Plumbley, L. K. Hansen, and S. H. Jensen, *Theorems on positive data: On the uniqueness of NMF*, Comput. Intell. Neurosci., 2008 (2008), 764206.

[22] R. Laursen and A. Hobolth, *A sampling algorithm to compute the set of feasible solutions for non-negative matrix factorization with an arbitrary rank*, SIAM J. Matrix Anal. Appl., 43 (2022), pp. 257–273.

[23] W. H. Lawton and E. A. Sylvestre, *Self modeling curve resolution*, Technometrics, 13 (1971), pp. 617–633.

[24] V. Leplat, N. Gillis, and J. Idier, *Multiplicative updates for NMF with β-divergences under disjoint equality constraints*, SIAM J. Matrix Anal. Appl., 42 (2021), pp. 730–752.

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

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

[27] E. R. Malinowski, *Factor Analysis in Chemistry*, 3rd ed., John Wiley and Sons, Hoboken, NJ, 2002.

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

[29] K. Neymeyr and M. Sawall, *On the set of solutions of the nonnegative matrix factorization problem*, SIAM J. Matrix Anal. Appl., 39 (2019), pp. 1049–1069.

[30] 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. Chemometr., 32 (2018), e2975.

[31] N. Omidikia, M. Ghaffari, and R. Rajkó, *Sparse non-negative multivariate curve resolution: L0, L1, or L2 norms?*, Chemometr. Intell. Lab. Syst., 199 (2020), 103969.

[32] R. Rajkó, *Natural duality in minimal constrained self modeling curve resolution*, J. Chemometr., 20 (2006), pp. 164–169.

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

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

[35] R. Rajkó, H. Abdollahi, S. Beyramysoltan, and N. Omidikia, *Definition and detection of data-based uniqueness in evaluating bilinear (two-way) chemical measurements*, Analy. Chim. Acta, 855 (2015), pp. 21–33.

[36] R. Rajkó and K. István, *Analytical solution for determining feasible regions of self-modeling curve resolution (SMCR) method based on computational geometry*, J. Chemometr., 19 (2005), pp. 448–463.

[37] M. Sawall, A. Jürss, H. Schröder, and K. Neymeyr, *On the analysis and computation of the area of feasible solutions for two-, three-, and four-component systems*, in Resolving Spectral Mixtures, Data Handling in Sci. Technol. 30, C. Ruckebusch, ed., Elsevier, New York, 2016, pp. 135–184.

[38] M. Sawall, C. Kubis, D. Selent, A. B. orner, and K. Neymeyr, *A fast polygon inflation algorithm to compute the area of feasible solutions for three-component systems.* I*: Concepts and applications*, J. Chemometr., 27 (2013), pp. 106–116.

[39] S. A. Vavasis, *On the complexity of nonnegative matrix factorization*, SIAM J. Optim., 20 (2010), pp. 1364–1377.

[40] M. Vosough, C. Mason, R. Tauler, M. Jalali-Heravi, and M. Maeder, *On rotational ambiguity in model-free analyses of multivariate data*, J. Chemometr., 20 (2006), pp. 302–310.

[41] R. Wehrens, *Chemometrics with R: Multivariate Data Analysis in* the *Natural and Life Sciences (Use R!)*, 2nd ed., Springer, Berlin, 2020.