



**HAL**  
open science

# Identifiability in Exact Multilayer Sparse Matrix Factorization

Léon Zheng, Rémi Gribonval, Elisa Riccietti

► **To cite this version:**

Léon Zheng, Rémi Gribonval, Elisa Riccietti. Identifiability in Exact Multilayer Sparse Matrix Factorization. 2021. hal-03362626v1

**HAL Id: hal-03362626**

**<https://inria.hal.science/hal-03362626v1>**

Preprint submitted on 1 Oct 2021 (v1), last revised 7 Oct 2022 (v6)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Identifiability in Exact Multilayer Sparse Matrix Factorization\*

Léon Zheng<sup>†</sup>, Rémi Gribonval<sup>†</sup>, and Elisa Riccietti<sup>†</sup>

**Abstract.** Many well-known matrices  $\mathbf{Z}$  are associated to fast transforms corresponding to factorizations of the form  $\mathbf{Z} = \mathbf{X}^{(L)} \dots \mathbf{X}^{(1)}$ , where each factor  $\mathbf{X}^{(\ell)}$  is sparse. Based on general result for the case with two factors, established in a companion paper, we investigate essential uniqueness of such factorizations. We show some identifiability results for the sparse factorization into two factors of the discrete Fourier Transform, discrete cosine transform or discrete sine transform matrices of size  $N = 2^L$ , when enforcing  $\frac{N}{2}$ -sparsity by column on the left factor, and 2-sparsity by row on the right factor. We also show that the analysis with two factors can be extended to the multilayer case, based on a hierarchical factorization method. We prove that any matrix which is the product of  $L$  factors whose supports are exactly the so-called butterfly supports, admits a unique sparse factorization into  $L$  factors. This applies in particular to the Hadamard or the discrete Fourier transform matrix of size  $2^L$ .

**Key words.** Identifiability, matrix factorization, sparsity, multilinear inverse problem, matrix completion

**AMS subject classifications.** 15A23, 15A83, 65F50, 94A12

**1. Introduction.** In order to understand under which conditions the problem of sparse matrix factorization is well-posed, we presented in the companion paper [14] a general framework to study identifiability, *i.e.* uniqueness, of exact sparse matrix factorization into two factors. The framework considered general family of sparsity patterns to encode sparsity constraints for matrix factorization. We took into account both scaling and permutation ambiguities which are inherent to classical families of sparsity patterns. We applied the lifting procedure [3] to establish equivalence between uniqueness of exact sparse matrix factorization, and uniqueness of exact sparse matrix decomposition into rank-one matrices, up to some non-degeneration properties on the factors.

In this paper, we apply this general framework to show some identifiability results on the sparse factorization of some well-known matrices, like the Discrete Fourier Transform (DFT) matrix, the Discrete Cosine Transform (DCT) matrix, or the Discrete Sine Transform (DST) matrix. The typical identifiability statement given in this paper takes the following form.

*Informal Theorem 1.1.* Let  $\mathbf{Z}$  be a matrix that we want to factorize, and  $\Omega$  be a *family of sparsity patterns*. If a certain condition on  $\Omega$  is satisfied, then  $\mathbf{Z}$  admits a unique sparse factorization  $\mathbf{Z} = \mathbf{X}\mathbf{Y}^\top$  up to some *unavoidable equivalences*, when the sparsity constraints on the factors are described by  $\Omega$ .

Examples of such a statement applied to the matrices mentioned above are given in [Theorem 2.25](#) or [Corollary 3.9](#). In complement to these results, we provide some sufficient conditions for identifiability of the two-layer exact sparse matrix factorization problem, when the sparsity constraint is encoded by a *fixed* pair of supports. These sufficient conditions are based on

---

\*This project was supported in part by the AllegroAssai ANR project ANR-19-CHIA-0009.

<sup>†</sup>Université de Lyon, ENS de Lyon, UCBL, CNRS, Inria, LIP, F-69342, LYON Cedex 07, France. ([leon.zheng@ens-lyon.fr](mailto:leon.zheng@ens-lyon.fr), [remi.gribonval@inria.fr](mailto:remi.gribonval@inria.fr), [elisa.riccietti@ens-lyon.fr](mailto:elisa.riccietti@ens-lyon.fr)).

rank-one matrix completability. We derive from these conditions an algorithm for the sparse reconstruction of the two factors when enforcing the fixed-support constraint.

This paper also shows how the framework of [14] for studying identifiability of sparse matrix factorization with two factors can be extended to the multilayer case with a hierarchical factorization method [11]. When enforcing the so-called butterfly supports (Definition 3.4) as the sparsity constraint for the multiple factors, we show in Corollary 3.9 that the DFT matrix admits a unique “butterfly” factorization [5], up to some rescaling equivalences. This identifiability results can be generalized to any matrix which is the product of factors whose supports are exactly the butterfly supports (Theorem 3.8).

The paper is organized as follows: section 2 recalls the framework introduced in the companion paper [14] to study identifiability in exact two-layer sparse matrix factorization, and shows an application of this framework for the sparse matrix factorization of DFT, DCT and DST matrices into two factors; section 3 is devoted to the extension of our analysis with two factors to the multilayer case; section 4 discusses perspectives of this work for algorithmic methods. An annex gathers technical lemmas and their proofs.

**2. Identifiability in two-layer sparse matrix factorization.** We first recall the framework to analyze identifiability in exact sparse matrix factorization into two factors introduced in the companion paper [14], with its main results. Subsection 2.2 is a word-to-word restatement of main results from the companion paper [14], and can be skipped by a familiar reader. Then, as new contributions of this paper, we provide in subsection 2.3 some sufficient conditions for fixed support identifiability, and we conclude in subsection 2.4 with some applications of the framework to establish identifiability results for the sparse factorization of the DFT, DCT and DST matrices.

**2.1. Notations.** The set of integers  $\{1, \dots, n\}$  is denoted  $\llbracket n \rrbracket$ . The cardinality of any finite set  $F$  is denoted  $\text{card}(F)$ . The complement of a set  $I$  is denoted  $I^c$ . The *support* of a matrix  $\mathbf{M} \in \mathbb{C}^{m \times n}$  of size  $m \times n$  is the set of indices  $\text{supp}(\mathbf{M}) \subseteq \llbracket m \rrbracket \times \llbracket n \rrbracket$  of nonzero entries. It is identified by abuse to the set of binary matrices  $\mathbb{B}^{m \times n} := \{0, 1\}^{m \times n}$ . Depending on the context, a matrix support can be seen as a set of indices, or a binary matrix with only nonzero entries for indices in this set. The cardinal of the matrix support  $\text{supp}(\mathbf{M})$  is also known as the  $\ell_0$ -norm of  $\mathbf{M}$ , denoted  $\|\cdot\|_0$ . The *column support*, denoted  $\text{colsupp}(\mathbf{M})$ , is the subset of indices  $i \in \llbracket n \rrbracket$  such that the  $i$ -th column of  $\mathbf{M}$ , denoted  $\mathbf{M}_i$ , is nonzero. Note that the notation  $\mathbf{M}_i$  with a bold subscript is reserved for a matrix, and should be distinguished from  $M_i$  which is a column of a matrix. Similarly, the *row support* of a matrix  $\mathbf{M}$  is the column support of its transpose, and is denoted  $\text{rowsupp}(\mathbf{M})$ . In other words, it is the subset of indices  $j \in \llbracket m \rrbracket$  such that the  $j$ -th row of  $\mathbf{M}$  is nonzero. The entry of  $\mathbf{M}$  indexed by  $(k, l)$  is  $M_{k,l}$ . The identity matrix of size  $n$  denoted  $\mathbf{I}_n$ . The Kronecker product between  $\mathbf{A}$  and  $\mathbf{B}$  is written  $\mathbf{A} \otimes \mathbf{B}$ . The set of permutations on a set of indices  $I$  is written  $\mathfrak{S}(I)$ . The function  $f$  iterated  $n$  times is denoted  $f^{(n)}$ .

**2.2. A general framework for the analysis of identifiability.** Given an observed matrix  $\mathbf{Z} \in \mathbb{C}^{m \times n}$ , and a subset of feasible pairs of factors  $\Sigma \subseteq \mathbb{C}^{m \times r} \times \mathbb{C}^{n \times r}$ , the so-called *exact matrix factorization* (EMF) problem with two factors of  $\mathbf{Z}$  in  $\Sigma$  is the following bilinear inverse

problem:

$$(EMF) \quad \text{find if possible } (\mathbf{X}, \mathbf{Y}) \in \Sigma \text{ such that } \mathbf{Z} = \mathbf{X}\mathbf{Y}^\top.$$

We are interested in the particular problem variation where the constraint set  $\Sigma$  encodes some chosen sparsity patterns for the factorization. For a given matrix support  $\mathbf{S} \in \mathbb{B}^{m \times r}$ , denote

$$(2.1) \quad \Sigma_{\mathbf{S}} := \{\mathbf{M} \in \mathbb{C}^{m \times r} \mid \text{supp}(\mathbf{M}) \subseteq \text{supp}(\mathbf{S})\}$$

the so-called *model-set* defined by  $\mathbf{S}$ , which is the set of matrices with a sparsity pattern included in  $\mathbf{S}$ . A pair of sparsity patterns is written  $\mathbf{S} := (\mathbf{S}^L, \mathbf{S}^R)$ , where  $\mathbf{S}^L$  and  $\mathbf{S}^R$  are the left and right sparsity patterns respectively. By abuse, we will refer to  $\mathbf{S}^L$  and  $\mathbf{S}^R$  as left and right (allowed) supports. Given any pair of allowed supports represented by binary matrices  $(\mathbf{S}^L, \mathbf{S}^R) \in \mathbb{B}^{m \times r} \times \mathbb{B}^{n \times r}$ , the set  $\Sigma_{\mathbf{S}^L} \times \Sigma_{\mathbf{S}^R} \subseteq \mathbb{C}^{m \times r} \times \mathbb{C}^{n \times r}$  is a linear subspace. Given any family  $\Omega \subset \mathbb{B}^{m \times r} \times \mathbb{B}^{n \times r}$  of such pairs of allowed supports, denote

$$(2.2) \quad \Sigma_{\Omega} := \bigcup_{\mathbf{S} \in \Omega} \Sigma_{\mathbf{S}}, \quad \text{with } \mathbf{S} := (\mathbf{S}^L, \mathbf{S}^R), \text{ and } \Sigma_{\mathbf{S}} := \Sigma_{\mathbf{S}^L} \times \Sigma_{\mathbf{S}^R}.$$

Such a set is a union of subspaces. Moreover, since the support of a matrix is unchanged under arbitrary rescaling of its columns,  $\Sigma_{\Omega}$  is invariant by column scaling for any family  $\Omega$ .

Our framework covers some classical families  $\Omega$  of structured sparse supports, like global sparsity, sparsity by column and/or row. In the following, the set of matrices (of a given size) which are  $k$ -sparse by column is denoted  $\Lambda_{\text{col}}^k$ :

$$(2.3) \quad \Lambda_{\text{col}}^k := \{\mathbf{M} \mid \forall j, \|\mathbf{M}_j\|_0 \leq k\},$$

These classical families are also invariant to permutations of columns, and will be referred to as families *stable* by permutation. Hence, uniqueness of a solution to (EMF) with such sparsity constraints will always be considered up to scaling and permutation equivalences. Denote  $\mathcal{P}_r$  as the group of permutation matrices of size  $r \times r$ ,  $\mathcal{D}_r$  as the group of diagonal matrices of size  $r \times r$  with nonzero diagonal entries, and  $\mathcal{G}_r := \{\mathbf{D}\mathbf{P} \mid \mathbf{D} \in \mathcal{D}_r, \mathbf{P} \in \mathcal{P}_r\}$  as the group of *generalized permutation* matrices of size  $r \times r$ .

**Definition 2.1 (PS-uniqueness of an EMF in  $\Sigma$  [14, Definition 2.5]).** *For any set  $\Sigma$  of pairs of factors, the pair  $(\mathbf{X}, \mathbf{Y}) \in \Sigma$  is the PS-unique EMF of  $\mathbf{Z} := \mathbf{X}\mathbf{Y}^\top$  in  $\Sigma$ , if any solution  $(\mathbf{X}', \mathbf{Y}')$  to (EMF) of  $\mathbf{Z}$  in  $\Sigma$  is equivalent to  $(\mathbf{X}, \mathbf{Y})$ , written  $(\mathbf{X}', \mathbf{Y}') \sim (\mathbf{X}, \mathbf{Y})$ , in the sense that there exists a permutation matrix  $\mathbf{P} \in \mathcal{P}_r$  and a diagonal matrix  $\mathbf{D} \in \mathcal{D}_r$  with nonzero diagonal entries such that  $(\mathbf{X}', \mathbf{Y}') = (\mathbf{X}\mathbf{D}\mathbf{P}, \mathbf{Y}\mathbf{D}^{-1}\mathbf{P})$ ; or, alternatively, there exists a generalized permutation matrix  $\mathbf{G} \in \mathcal{G}_r$  such that  $(\mathbf{X}', \mathbf{Y}') = (\mathbf{X}\mathbf{G}, \mathbf{Y}(\mathbf{G}^{-1})^\top)$ .*

For any set  $\Sigma$  of pairs of factors, the set of all pairs  $(\mathbf{X}, \mathbf{Y}) \in \Sigma$  such that  $(\mathbf{X}, \mathbf{Y})$  is the PS-unique EMF of  $\mathbf{Z} := \mathbf{X}\mathbf{Y}^\top$  in  $\Sigma$  is denoted  $\mathcal{U}(\Sigma)$ . Alternatively, we define:

$$(2.4) \quad \mathcal{U}(\Sigma) := \{(\mathbf{X}, \mathbf{Y}) \in \Sigma \mid \forall (\mathbf{X}', \mathbf{Y}') \in \Sigma, \mathbf{X}'\mathbf{Y}'^\top = \mathbf{X}\mathbf{Y}^\top \implies (\mathbf{X}', \mathbf{Y}') \sim (\mathbf{X}, \mathbf{Y})\}.$$

In order to characterize the set  $\mathcal{U}(\Sigma_{\Omega})$  for  $\Sigma_{\Omega}$  defined as in (2.2) with  $\Omega$  any family that is *stable* by permutation, two non-degeneration properties were established for a pair of factors

[14], i.e. two necessary conditions for identifiability, involving their so-called *column support*. The first non-degeneration property for identifiability of EMF in  $\Sigma_\Omega$  thus requires the left and right factors to have *the same column supports*. Define the set of pairs of factors with *identical column supports* in  $\Sigma_\Omega$  as

$$(2.5) \quad \text{IC}_\Omega := \{(\mathbf{X}, \mathbf{Y}) \in \Sigma_\Omega \mid \text{colsupp}(\mathbf{X}) = \text{colsupp}(\mathbf{Y})\}.$$

The second non-degeneration property for a pair of factors requires the column supports of the left and right factors to be “maximal”. Define the set of pairs of factors with *maximal column supports* in  $\Sigma_\Omega$  as

$$(2.6) \quad \text{MC}_\Omega := \{(\mathbf{X}, \mathbf{Y}) \in \Sigma_\Omega \mid \forall \mathbf{S} \in \Omega \text{ such that } (\mathbf{X}, \mathbf{Y}) \in \Sigma_{\mathbf{S}}, \text{colsupp}(\mathbf{X}) = \text{colsupp}(\mathbf{S}^L) \\ \text{and } \text{colsupp}(\mathbf{Y}) = \text{colsupp}(\mathbf{S}^R)\}.$$

**Lemma 2.2 (From [14, Lemma 2.11]).** *For any family of pairs of supports  $\Omega$ , we have:  $\mathcal{U}(\Sigma_\Omega) \subseteq \text{IC}_\Omega \cap \text{MC}_\Omega$ .*

As the matrix product  $\mathbf{X}\mathbf{Y}^\top$  can be decomposed into the sum of rank-one matrices  $\sum_{i=1}^r \mathbf{X}_i \mathbf{Y}_i^\top$ , the so-called lifting procedure [3, 13] suggests to represent a pair  $(\mathbf{X}, \mathbf{Y})$  by its  $r$ -tuple of so-called *rank-one contributions*. With this lifting approach, each support constraint  $\mathbf{S} = (\mathbf{S}^L, \mathbf{S}^R)$  is represented by the  $r$ -tuple of rank-one support constraints  $\mathcal{S} = \varphi(\mathbf{S}^L, \mathbf{S}^R)$ . Thus, when  $(\mathbf{X}, \mathbf{Y})$  follows a sparsity structure given by  $\Omega$ , i.e. belongs to  $\Sigma_\Omega \subseteq \mathbb{C}^{m \times r} \times \mathbb{C}^{n \times r}$ , the  $r$ -tuple of rank-one matrices  $\varphi(\mathbf{X}, \mathbf{Y}) \in (\mathbb{C}^{m \times n})^r$  belongs to the set:

$$(2.7) \quad \Gamma_\Omega := \bigcup_{\mathcal{S} \in \varphi(\Omega)} \Gamma_{\mathcal{S}}, \quad \text{with } \mathcal{S} := (\mathbf{S}^1, \dots, \mathbf{S}^r), \quad \text{and}$$

$$(2.8) \quad \Gamma_{\mathcal{S}} := \{(\mathbf{C}^i)_{i=1}^r \mid \forall i \in \llbracket r \rrbracket, \text{rank}(\mathbf{C}^i) \leq 1, \text{supp}(\mathbf{C}^i) \subseteq \mathbf{S}^i\} \subseteq (\mathbb{C}^{m \times n})^r.$$

The most important property of the lifting approach with respect to identifiability is that PS-uniqueness of an EMF in  $\Sigma = \text{IC}_\Omega$  is equivalent to identifiability of the rank-one contributions in  $\Gamma_\Omega$ . Denote

$$(2.9) \quad \mathcal{A} : \mathcal{C} = (\mathbf{C}^i)_{i=1}^r \mapsto \sum_{i=1}^r \mathbf{C}^i$$

the linear operator which sums the  $r$  matrices of a tuple  $\mathcal{C}$ .

**Definition 2.3 (P-uniqueness of an EMD in  $\Gamma$  [14, Definition 4.5]).** *For any set  $\Gamma \subseteq (\mathbb{C}^{m \times n})^r$  of  $r$ -tuples of rank-one matrices, the  $r$ -tuple  $\mathcal{C} \in \Gamma$  is the P-unique exact matrix decomposition (EMD) of  $\mathbf{Z} := \mathcal{A}(\mathcal{C})$  in  $\Gamma$  if, for any  $\mathcal{C}' \in \Gamma$  such that  $\mathcal{A}(\mathcal{C}') = \mathbf{Z}$ , we have  $\mathcal{C}' \sim \mathcal{C}$ , in the sense that the tuples  $(\mathbf{C}^i)_{i=1}^r$  and  $(\mathbf{C}'^i)_{i=1}^r$  are equal up to a permutation of the index  $i$ .*

The set of all  $r$ -tuples  $\mathcal{C} \in \Gamma$  such that  $\mathcal{C}$  is the P-unique EMD of  $\mathbf{Z} := \mathcal{A}(\mathcal{C})$  in  $\Gamma$  is denoted  $\mathcal{U}(\Gamma)$ , where the notation  $\mathcal{U}(\cdot)$  has been slightly abused, as  $\Gamma$  and  $\Sigma$  are subsets of a different nature.

**Theorem 2.4 (From [14, Theorem 4.8]).** *For any family of pairs of supports  $\Omega$  stable by permutation, and any pair of factors  $(\mathbf{X}, \mathbf{Y})$ :*

$$(\mathbf{X}, \mathbf{Y}) \in \mathcal{U}(\Sigma_\Omega) \iff \varphi(\mathbf{X}, \mathbf{Y}) \in \mathcal{U}(\Gamma_\Omega) \text{ and } (\mathbf{X}, \mathbf{Y}) \in \text{IC}_\Omega \cap \text{MC}_\Omega.$$

Hence, it is sufficient to characterize  $(\mathbf{X}, \mathbf{Y}) \in \text{IC}_\Omega \cap \text{MC}_\Omega$  such that the  $r$ -tuple of rank-one contributions  $\varphi(\mathbf{X}, \mathbf{Y})$  is the P-unique EMD of  $\mathbf{Z} := \mathbf{X}\mathbf{Y}^\top$  in  $\Gamma_\Omega$ . Our approach to characterize P-uniqueness of an EMD relied on a strategy based on the following result.

**Proposition 2.5** (From [14, Proposition 4.9]). *For any family of pairs of supports  $\Omega$  stable by permutation, we have  $\mathcal{C} \in \mathcal{U}(\Gamma_\Omega)$  if, and only if, both of the following conditions are verified:*

- (i) *For all  $\mathcal{S} \in \varphi(\Omega)$  such that  $\mathcal{A}(\mathcal{C}) \in \mathcal{A}(\Gamma_\mathcal{S})$ , there exists a permutation  $\sigma \in \mathfrak{S}(\llbracket r \rrbracket)$  such that:  $\text{supp}(\mathbf{C}^i) \subseteq \mathcal{S}^{\sigma(i)}$  for each  $1 \leq i \leq r$ .*
- (ii) *For all  $\mathcal{S} \in \varphi(\Omega)$  such that  $\mathcal{C} \in \Gamma_\mathcal{S}$ , we have  $\mathcal{C} \in \mathcal{U}(\Gamma_\mathcal{S})$ .*

A simple sufficient condition for condition (i) of the previous proposition can be established [14]. Denote:

$$(2.10) \quad \varphi^{\text{cp}}(\Omega) := \bigcup_{\mathcal{S} \in \varphi(\Omega)} \left\{ \mathcal{S}' \in (\mathbb{B}^{m \times n})^r \mid \forall i = 1, \dots, r, \text{rank}(\mathbf{S}'^i) \leq 1, \mathcal{S}'^i \subseteq \mathcal{S}^i \right\}.$$

We recall that the inclusion  $\mathcal{S}'^i \subseteq \mathcal{S}^i$  means the inclusion of the supports viewed as subsets of indices.

**Proposition 2.6** (From [14, Proposition 4.11]). *Consider  $\mathcal{C} \in \Gamma_\Omega$  and  $\mathbf{Z} := \mathcal{A}(\mathcal{C})$ , and assume that:*

- (i) *for each  $\mathcal{C}' \in \Gamma_\Omega$  such that  $\mathcal{A}(\mathcal{C}') = \mathbf{Z}$ , the supports  $\{\text{supp}(\mathbf{C}'^i)\}_{i=1}^r$  are pairwise disjoint;*
- (ii) *all  $\mathcal{S} \in \varphi^{\text{cp}}(\Omega)$  such that  $\{\mathcal{S}^i\}_{i=1}^r$  is a partition of  $\text{supp}(\mathbf{Z})$ , and such that the rank of  $(\mathbf{Z}_{k,l})_{(k,l) \in \mathcal{S}^i}$  ( $1 \leq i \leq r$ ) is at most one, are equivalent up to a permutation.*

*Then, the supports  $\{\text{supp}(\mathbf{C}^i)\}_{i=1}^r$  are identifiable in the sense of condition (i) in Proposition 2.5.*

Our first main contribution, in subsection 2.4, is to show that these sufficient conditions are verified in practice to give some identifiability results on the DFT, DCT and DST matrices. In subsection 2.3, we show some sufficient conditions for identifiability in exact matrix decomposition with fixed supports, as a complement of some necessary conditions given in the companion paper [14], which were based on an analysis of identifiability of the right factor when the left one is fixed. We also derive a new algorithm based on rank-one matrix completion for the exact recovery of the two sparse factors, in the case of the fixed-support constraint.

**2.3. Sufficient conditions for fixed support identifiability.** The sufficient condition for fixed-support identifiability in EMD given in this section is based on a notion of *closability* in some bipartite graphs associated to “observable entries” in rank-one supports.

**2.3.1. Observable vs missing entries, bipartite graphs, and matrix completion.** Given an  $r$ -tuple  $\mathcal{S}$  of rank-one supports and  $\mathcal{C} \in \Gamma_\mathcal{S}$ , for any  $i \in \llbracket r \rrbracket$ , an entry is said to be *missing* in  $\mathbf{C}^i$  if its index  $(k, l)$  belongs to the intersection  $\mathcal{S}^i \cap \mathcal{S}^j$  for another  $j \neq i$ ; otherwise, it is said to be *observable*. The intuition behind these notions is that each observable entry is equal to the corresponding entry of the factorized matrix  $\sum_i \mathbf{C}^i$ , while missing entries need to be deduced, if possible, from observable ones using the low-rank constraints.

Low-rank matrix completion problems without sparsity constraints are naturally associated to properties of certain bipartite graphs [7, Proposition 2.15]. This fact was exploited to prove fixed-support identifiability of certain factorizations of the DFT [10, Chapter 7] using completion operations inside each of the rank-one contributions. More precisely, it was established that when all the missing values inside each contribution can be completed without ambiguity from the observable ones using the rank-one constraint, the considered tuple of rank-one contributions is identifiable for the EMD with fixed support. We extend these results by showing a more refined sufficient condition for  $\mathcal{C} \in \mathcal{U}(\Gamma_{\mathcal{S}})$ : all the missing entries in the contributions  $\{\mathcal{C}^i\}_{i=1}^r$  can be recovered through *iterative* and *partial* completion operations based on the rank-one constraint, starting only from the knowledge of the observable entries. These completion operations include completion inside each contribution, as well completion *across* the contributions.

**Notations for graphs.** A graph  $\mathcal{G}$  is defined by a set of vertices  $V$  and a set of edges  $E \subseteq V^2$ , and is denoted  $\mathcal{G}(V, E)$ . The set of vertices and the set of edges in a given graph  $\mathcal{G}$  are denoted respectively  $V(\mathcal{G})$  and  $E(\mathcal{G})$ . The complement of a graph  $\mathcal{G}(V, E)$  is the graph  $\overline{\mathcal{G}} = \mathcal{G}(V, V^2 \setminus E)$ . In a *bipartite* graph  $\mathcal{G}$ , the set of vertices is partitioned into two sets of vertices  $V_1$  and  $V_2$ , and its set of edges satisfies  $E \subseteq V_1 \times V_2$ . This will be denoted  $\mathcal{G}(V_1, V_2, E)$ . By convention, the elements in the first group of vertices  $V_1$  will be called red vertices, while the elements in the second group of vertices  $V_2$  will be called blue vertices. Then, the set of red vertices and the set of blue vertices for a given bipartite graph  $\mathcal{G}$  are denoted by  $V_1(\mathcal{G})$  and  $V_2(\mathcal{G})$  respectively. A bipartite graph  $\mathcal{G}(V_1, V_2, E)$  with  $V_1 \subseteq \llbracket n \rrbracket$ ,  $V_2 \subseteq \llbracket m \rrbracket$  is a representation of a matrix support  $\mathcal{S}$  of size  $n \times m$ , where the nonzero entries corresponds to the set  $E$ . In other words,  $\mathcal{S}$  is the adjacency matrix of  $\mathcal{G}$ . This justifies why we use bold notation for a graph, even though it is not a matrix. The *complete bipartite graph*  $\mathcal{G}(V_1, V_2, V_1 \times V_2)$  is denoted  $\mathcal{K}_{V_1, V_2}$ . Given a bipartite graph  $\mathcal{G}$ , we denote  $\mathcal{K}_{\mathcal{G}}$  the corresponding *completed bipartite graph* defined as  $\mathcal{K}_{\mathcal{G}} := \mathcal{K}_{V_1(\mathcal{G}), V_2(\mathcal{G})}$ . An  $r$ -tuple of graphs is denoted  $\mathcal{G} := (\mathcal{G}^1, \dots, \mathcal{G}^r)$ . In the following, we consider only undirected graphs.

We define bipartite graphs whose adjacency matrices correspond to observable entries.

**Definition 2.7 (Observable supports, observable bipartite graphs).** *Consider an  $r$ -tuple of rank-one supports  $\mathcal{S} = (\mathcal{S}^1, \dots, \mathcal{S}^r)$ . Viewing each  $\mathcal{S}^i$  as an index set, the observable supports are the subsets:*

$$(2.11) \quad \text{obs}_i(\mathcal{S}) := \mathcal{S}^i \setminus \left( \bigcup_{j \neq i} \mathcal{S}^j \right), \quad 1 \leq i \leq r.$$

*Viewing each  $\mathcal{S}^i$  as a binary matrix of rank one, the  $i$ -th observable bipartite graph of  $\mathcal{S}$  ( $1 \leq i \leq r$ ) is the bipartite graph  $\mathcal{G}^i(\mathcal{S}) := (\text{rowsupp}(\mathcal{S}^i), \text{colsupp}(\mathcal{S}^i), \text{obs}_i(\mathcal{S}))$ . The  $r$ -tuple of observable bipartite graphs  $\mathcal{G}^i(\mathcal{S})$ ,  $1 \leq i \leq r$ , associated to  $\mathcal{S}$  is denoted  $\mathcal{G}(\mathcal{S})$ .*

This is illustrated on an example in [Figure 1](#).

The role of bipartite graphs in characterizing fixed-support identifiability in EMD will become apparent once we recall an existing result in low-rank matrix completion. For any matrix support  $\mathcal{S} \in \mathbb{B}^{m \times n}$  interpreted as a binary mask, and any observed submatrix  $(\mathbf{Z}_{k,l})_{(k,l) \in \mathcal{S}}$ , define the *rank-one matrix completion* problem as:

$$(2.12) \quad \text{find, if possible, } \mathbf{M} \text{ of rank at most one such that } (\mathbf{M}_{ij})_{(i,j) \in \mathcal{S}} = (\mathbf{Z}_{k,l})_{(k,l) \in \mathcal{S}}.$$

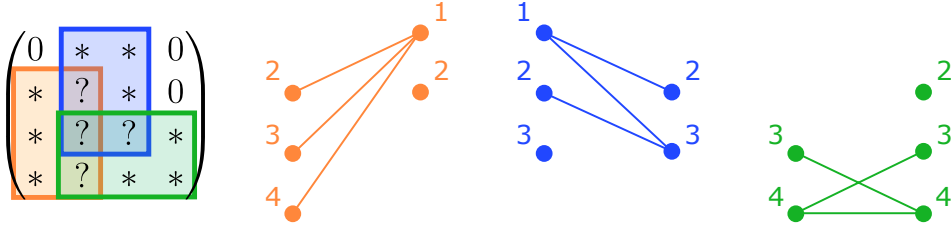


Figure 1: Illustration of the observable bipartite graphs of  $\mathcal{S} := (\mathcal{S}^1, \mathcal{S}^2, \mathcal{S}^3)$  where  $\mathcal{S}^1 = \{2, 3, 4\} \times \{1, 2\}$  (orange),  $\mathcal{S}^2 = \{1, 2, 3\} \times \{2, 3\}$  (blue) and  $\mathcal{S}^3 = \{3, 4\} \times \{2, 3, 4\}$  (green). The indices marked with “\*” are those which belong to only one rank-one support (observable supports). The indices marked with “?” are those which belong to at least two different rank-one supports.

The next proposition from [7], illustrated by Figure 2, characterizes uniqueness in rank-one matrix completion.

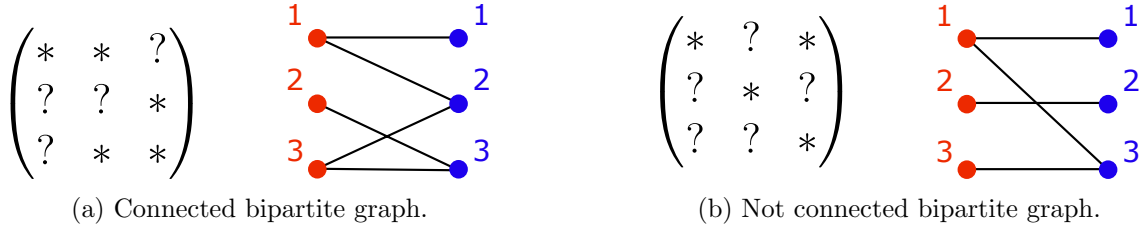


Figure 2: Illustration of Proposition 2.8: on the left, it is possible to complete the missing values denoted by “?” from the nonzero observed values “\*” in the mask, because the bipartite graph associated to this mask is connected. On the right, this condition is not verified.

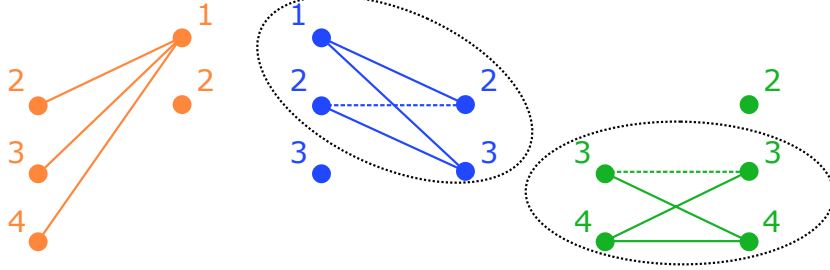
**Proposition 2.8 (Adapted from [7, Proposition 2.15]).** *A rank-one matrix  $\mathbf{A}$  of size  $m \times n$  which has no zero entry is the unique solution of the completion problem (2.12) for the observed submatrix  $(\mathbf{A}_{k,l})_{(k,l) \in \mathcal{S}}$  on the mask  $\mathcal{S}$  if, and only if, the bipartite graph  $\mathcal{G}(\llbracket m \rrbracket, \llbracket n \rrbracket, \mathcal{S})$  is connected.*

**2.3.2. Completion operations for tuples of bipartite graphs.** The sufficient conditions for fixed-support identifiability that we will express rely on two completion operations in the observable bipartite graphs  $\mathcal{G}(\mathcal{S})$  that we now define. These operations are illustrated in Figure 3 for the tuple of graphs introduced in Figure 1.

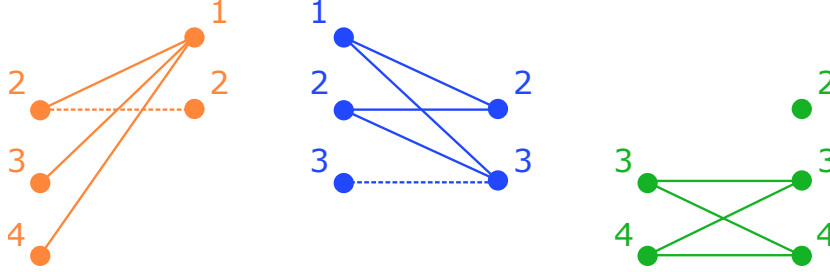
The first completion operation is directly inspired from Proposition 2.8 and [10, Chapter 7]. It corresponds to the completion of all missing edges in all connected subgraphs in each bipartite graph. In terms of completion of missing values in rank-one matrices, it corresponds to solving successively trivial linear systems with one variable. For instance, knowing that  $(a \ x)^T$  and  $(b \ c)$  are collinear,  $x$  can be deduced from the values of  $a$ ,  $b$ , and  $c$ , as long as  $b$  is nonzero.



The second completion operation involves all the bipartite graphs in the tuple. In terms of completion of missing values in rank-one contributions, it corresponds to the fact that, for a given index pair  $(k, l)$ , when all the entries  $\mathcal{C}^j_{k,l}$  ( $j \in \llbracket r \rrbracket$ ) are known except for  $j = i$ , the entry  $\mathcal{C}^i_{k,l}$  is easily deduced from the knowledge of  $\mathbf{Z}_{k,l} := \sum_{j=1}^r \mathcal{C}^j_{k,l}$ .



(a) Completion *inside* each graph. We look for all the subgraphs isomorphic to  $\mathcal{K}_{2,2}^-$  (circled by dashed lines), and we complete the edge that is missing in each of these subgraphs.



(b) Completion *across* graphs. The edge  $(2, 2)$  belongs to the complete bipartite graphs  $\mathcal{K}_{\mathcal{G}^1}$  and  $\mathcal{K}_{\mathcal{G}^2}$ , and it is missing in  $\mathcal{G}^1$  but not in  $\mathcal{G}^2$ , so this edge is added to  $\mathcal{G}^1$  by the completion operation across graphs. Similarly, the edge  $(3, 3)$  is added to  $\mathcal{G}^2$ .

Figure 3: Illustration of completion operations defined by Definition 2.9. The tuple of bipartite graphs before completion  $(\mathcal{G}^1, \mathcal{G}^2, \mathcal{G}^3)$ , already introduced in Figure 1, is represented by full lines. Dashed lines represent edges that are added after the completion operation.

**Definition 2.9 (Completion operations).** Let  $\mathcal{G} := (\mathcal{G}^i)_{i=1}^r$  be an  $r$ -tuple of bipartite graphs. Denote  $\mathcal{K}_{2,2}^-$  the complete bipartite graph  $\mathcal{K}_{\llbracket 2 \rrbracket, \llbracket 2 \rrbracket}$  minus one edge.

- **Completion inside each graph:** For each  $i \in \llbracket r \rrbracket$ , define  $\mathcal{G}'^i$  as the bipartite graph which is obtained from  $\mathcal{G}^i$  by completing all the missing edges in each subgraph of  $\mathcal{G}^i$  that is isomorphic to  $\mathcal{K}_{2,2}^-$ , and denote  $a(\mathcal{G}) := (\mathcal{G}'^i)_{i=1}^r$ .
- **Completion across graphs:** If a given edge  $e$  is missing only in the graph  $\mathcal{G}^i$  but not in the other graphs  $\mathcal{G}^j$  for  $j \neq i$ , we complete this missing edge in  $\mathcal{G}^i$ . Formally, for each  $i \in \llbracket r \rrbracket$ , define  $\mathcal{G}''^i$  as the bipartite graph which is obtained from  $\mathcal{G}^i$  by adding all the edges in the set  $\left\{ e \in E(\mathcal{K}_{\mathcal{G}^i}) \mid e \notin E(\mathcal{G}^i) \text{ and } e \in \bigcap_{\substack{j \neq i \\ e \in E(\mathcal{K}_{\mathcal{G}^j})}} E(\mathcal{G}^j) \right\}$ .

*Remark 2.10.* In the spirit of [Proposition 2.8](#), the completion operation inside each graph can be defined equivalently in the following way: “For each  $i \in \llbracket r \rrbracket$ , define  $\mathcal{G}^i$  as the bipartite graph which is obtained from  $\mathcal{G}^i$  by completing all the missing edges in each connected subgraph of  $\mathcal{G}^i$ .”

The so-called *closure*  $\text{cl}(\mathcal{G})$  is then obtained by completing iteratively the missing edges in  $\mathcal{G}$  through the completion operations  $a$  and  $b$  of [Definition 2.9](#), until no more edges can be added. This process indeed reaches a fixed point after a finite number of steps.

**Lemma 2.11.** *Let  $\mathcal{G}$  be an  $r$ -tuple of bipartite graphs. There is a positive integer  $N$  such that  $(b \circ a)^{(N+1)}(\mathcal{G}) = (b \circ a)^{(N)}(\mathcal{G})$ , where  $a$  and  $b$  are the completion operations introduced in [Definition 2.9](#).*

The proof exploits a partial order between bipartite graphs sharing the same red and blue vertices, that will also be used in the proof of [Theorem 2.15](#) below.

**Definition 2.12.** *For any bipartite graphs  $\mathcal{G}, \mathcal{H}$  for which  $(V_1(\mathcal{G}), V_2(\mathcal{G})) = (V_1(\mathcal{H}), V_2(\mathcal{H}))$ , we write  $\mathcal{G} \preceq \mathcal{H}$  if  $E(\mathcal{G}) \subseteq E(\mathcal{H})$ . We extend this notation to  $r$ -tuples of bipartite graphs  $\mathcal{G}, \mathcal{H}$ :*

$$\mathcal{G} \preceq \mathcal{H} \iff \forall i \in \llbracket r \rrbracket, \mathcal{G}^i \preceq \mathcal{H}^i.$$

One verifies that this partial order is indeed reflexive, anti-symmetric and transitive.

*Proof of Lemma 2.11.* First, we remark that  $\mathcal{G} \preceq a(\mathcal{G}) \preceq (b \circ a)(\mathcal{G})$ . Secondly, for any bipartite graph  $\mathcal{H}$  with red vertices  $V_1$  and blue vertices  $V_2$ ,  $\mathcal{H} \preceq \mathcal{K}_{V_1, V_2}$ . But the number of edges in the complete bipartite graph  $\mathcal{K}_{V_1, V_2}$  is finite. In conclusion, we obtain the claimed result by applying the monotone convergence theorem.  $\blacksquare$

**Definition 2.13 (Closure of a tuple of bipartite graphs).** *The closure of an  $r$ -tuple of bipartite graphs  $\mathcal{G}$  is the  $r$ -tuple of bipartite graphs  $\text{cl}(\mathcal{G}) := (b \circ a)^{(N(\mathcal{G}))}(\mathcal{G})$  where*

$$(2.13) \quad N(\mathcal{G}) := \min \{N \in \mathbb{N} \mid (b \circ a)^{(N+1)}(\mathcal{G}) = (b \circ a)^{(N)}(\mathcal{G})\}.$$

**2.3.3. Sufficient condition for identifiability with fixed support.** We can now formulate our sufficient condition for identifiability with fixed support of  $\mathcal{C} \in \mathcal{U}(\Gamma_{\mathcal{S}})$ .

**Definition 2.14 (Closable tuple of rank-one supports).** *We say that the  $r$ -tuple of rank-one supports  $\mathcal{S}$  is closable if the closure of its observable bipartite graphs,  $\text{cl}(\mathcal{G}(\mathcal{S}))$ , is the  $r$ -tuple of complete bipartite graphs  $(\mathcal{K}_{\mathcal{G}^i(\mathcal{S})})_{i=1}^r$ .*

**Theorem 2.15.** *Suppose that  $\mathcal{S}$  is a closable  $r$ -tuple of rank-one supports. Denote the set of  $r$ -tuples of rank-one matrices with a support exactly equal to  $\mathcal{S}$  as:*

$$(2.14) \quad \Gamma_{\mathcal{S}}^{(=)} := \{\mathcal{C} \in \Gamma_{\mathcal{S}} \mid \forall i \in \llbracket r \rrbracket, \text{supp}(\mathcal{C}^i) = \mathcal{S}^i\}.$$

Any  $\mathcal{C} \in \Gamma_{\mathcal{S}}^{(=)}$  is the  $P$ -unique EMD of  $\mathbf{Z} := \mathcal{A}(\mathcal{C})$  in  $\Gamma_{\mathcal{S}}^{(=)}$ . In other words,  $\Gamma_{\mathcal{S}}^{(=)} \subseteq \mathcal{U}(\Gamma_{\mathcal{S}})$ .

The proof is deferred to [Appendix A](#).

It can be shown [[10](#), Chapter 7, Corollary 1] that any  $r$ -tuple of rank-one supports  $\mathcal{S}$  such that  $\mathcal{S}^i \cap \mathcal{S}^{i'} = \emptyset$  for all  $i \neq i'$  is closable. As an immediate corollary we obtain  $P$ -uniqueness of the corresponding rank-one factors.

**Corollary 2.16.** *Suppose that  $\mathcal{S}$  is such that  $\mathbf{S}^i \cap \mathbf{S}^{i'} = \emptyset$  for all  $i \neq i'$ . Then  $\Gamma_{\mathcal{S}}^{(=)} \subseteq \mathcal{U}(\Gamma_{\mathcal{S}})$ .*

As a matter of fact, the companion paper [14] showed that, in the case where the rank-one supports in  $\mathcal{S}$  are disjoint, *uniform P-uniqueness* of the EMD in  $\Gamma_{\mathcal{S}}$  holds, which is a stronger result than the previous corollary.

**Proposition 2.17.** *Let  $\mathcal{S}$  be an  $r$ -tuple of rank-one supports. Then, uniform P-uniqueness of EMD in  $\Gamma_{\mathcal{S}}$  holds, i.e.  $\mathcal{U}(\Gamma_{\mathcal{S}}) = \Gamma_{\mathcal{S}}$ , if, and only if, one of the following equivalent assertions holds:*

- (i) *The linear application  $\mathcal{A} : \mathcal{C} \mapsto \sum_{i=1}^r \mathcal{C}^i$  restricted to  $\Gamma_{\mathcal{S}}$  is injective.*
- (ii) *The rank-one supports  $\{\mathbf{S}^i\}_{i=1}^r$  are pairwise disjoint.*

Consequently, when a pair of supports  $\mathbf{S} := (\mathbf{S}^L, \mathbf{S}^R)$  is such that  $\varphi(\mathbf{S})$  has disjoint rank-one supports, almost every pair of factors  $(\mathbf{X}, \mathbf{Y})$  is identifiable for the EMF of  $\mathbf{Z} = \mathbf{X}\mathbf{Y}^\top$  in  $\Sigma_{\mathbf{S}}$ . In fact, a stronger identifiability than PS-uniqueness is verified in this case, which is identifiability up to scaling ambiguities *only*.

**Definition 2.18 (S-uniqueness in the EMF).** *For any set  $\Sigma$  of pairs of factors, the pair  $(\mathbf{X}, \mathbf{Y}) \in \Sigma$  is the S-unique EMF of  $\mathbf{Z} := \mathbf{X}\mathbf{Y}^\top$  in  $\Sigma$ , if any solution  $(\mathbf{X}', \mathbf{Y}')$  to (EMF) of  $\mathbf{Z}$  in  $\Sigma$  is equivalent to  $(\mathbf{X}, \mathbf{Y})$  up to scaling ambiguities only, written  $(\mathbf{X}', \mathbf{Y}') \sim_s (\mathbf{X}, \mathbf{Y})$*

The set of all pairs  $(\mathbf{X}, \mathbf{Y}) \in \Sigma$  such that  $(\mathbf{X}, \mathbf{Y})$  is the S-unique EMF of  $\mathbf{Z} := \mathbf{X}\mathbf{Y}^\top$  in  $\Sigma$  is denoted  $\mathcal{U}_s(\Sigma)$ .

**Corollary 2.19.** *Suppose that  $\mathbf{S} := (\mathbf{S}^L, \mathbf{S}^R)$  is such that  $\varphi(\mathbf{S})$  has disjoint rank-one supports. Then:*

$$\mathcal{U}_s(\Sigma_{\mathbf{S}}) = \text{IC}_{\mathbf{S}} \cap \text{MC}_{\mathbf{S}},$$

where  $\text{IC}_{\mathbf{S}}$  and  $\text{MC}_{\mathbf{S}}$  are the specialization of  $\text{IC}_{\Omega}$  and  $\text{MC}_{\Omega}$  defined by (2.5) and (2.6) to the case where  $\Omega$  is reduced to a singleton  $\mathbf{S}$ :

$$(2.15) \quad \text{IC}_{\mathbf{S}} := \{(\mathbf{X}, \mathbf{Y}) \in \Sigma_{\mathbf{S}} \mid \text{colsupp}(\mathbf{X}) = \text{colsupp}(\mathbf{Y})\},$$

$$(2.16) \quad \text{MC}_{\mathbf{S}} := \{(\mathbf{X}, \mathbf{Y}) \in \Sigma_{\mathbf{S}} \mid \text{colsupp}(\mathbf{X}) = \text{colsupp}(\mathbf{S}^L), \text{colsupp}(\mathbf{Y}) = \text{colsupp}(\mathbf{S}^R)\}.$$

However, a closable  $r$ -tuple of rank-one supports  $\mathcal{S}$  does not have necessarily pairwise disjoint rank-one supports. Consider for instance the tuple  $\mathcal{S}$  of Figure 1. Even though its rank-one supports are not pairwise disjoint,  $\mathcal{S}$  will also be shown to be closable (see discussion around Figure 4). The following example illustrates an application of Theorem 2.15 for a given  $\mathcal{C} \in \Gamma_{\mathcal{S}}$ .

**Example 2.20.** Denote  $\mathcal{S} := (\mathbf{S}^1, \mathbf{S}^2, \mathbf{S}^3)$  as the tuple of supports of size  $4 \times 4$  defined in Figure 1, where  $\mathbf{S}^1 = \{2, 3, 4\} \times \{1, 2\}$ ,  $\mathbf{S}^2 = \{1, 2, 3\} \times \{2, 3\}$ , and  $\mathbf{S}^3 = \{3, 4\} \times \{2, 3, 4\}$ . Define also:

$$\mathbf{Z} = \begin{pmatrix} 0 & 1 & 2 & 0 \\ 1 & 2 & 2 & 0 \\ 2 & 6 & 5 & 6 \\ 3 & 5 & 2 & 4 \end{pmatrix}.$$

Then, by [Theorem 2.15](#), the matrix  $\mathbf{Z}$  admits a P-unique EMD in  $\Gamma_{\mathcal{S}}$ , because any  $\mathcal{C} \in \Gamma_{\mathcal{S}}$  such that  $\mathcal{A}(\mathcal{C}) = \mathbf{Z}$  belongs to  $\Gamma_{\mathcal{S}}^{(=)}$ , and  $\mathcal{S}$  is closable.

**2.3.4. Algorithm based on rank-one matrix completion.** The closability condition of the previous section suggests an algorithm based on rank-one matrix completion to decompose, if possible, an observed matrix  $\mathbf{Z}$  into a sum of rank-one matrices  $\mathbf{C}^i$  ( $1 \leq i \leq r$ ) satisfying the rank-one sparsity constraints associated to  $\mathcal{S}$ . By design, [Algorithm 2.1](#) greedily completes missing values only if the completion is non ambiguous. On the one hand, in the case when there exist some ambiguities during the completion, the algorithm returns a tuple of rank-one contributions with remaining missing values. On the other hand, the absence of missing values of the output indicates that the input  $\mathbf{Z}$  admits a P-unique EMD with fixed supports constraint  $\mathcal{S}$ .

**Proposition 2.21.** *Let us run [Algorithm 2.1](#) with a given  $r$ -tuple of rank-one supports  $\mathcal{S}$  and a given observed matrix  $\mathbf{Z}$  as inputs. The following assertions hold:*

- (i) *If the algorithm breaks the loop because of incompatibility at [Line 15](#), then  $\mathbf{Z} \notin \mathcal{A}(\Gamma_{\mathcal{S}})$ .*
- (ii) *If the algorithm does not break the loop because of incompatibility, and outputs  $\mathcal{C}$  without missing values “?”, then  $\mathbf{Z} = \mathcal{A}(\mathcal{C})$ . Moreover,  $\mathcal{C}$  is the unique solution to  $\mathcal{A}(\mathcal{C}') = \mathbf{Z}$  such that  $\mathcal{C}' \in \Gamma_{\mathcal{S}}$ . A fortiori,  $\mathcal{C} \in \mathcal{U}(\Gamma_{\mathcal{S}})$ .*

*Proof.* To prove the first assertion, one can assume by contradiction that the algorithm breaks the loop because of incompatibility, with  $\mathbf{Z} \in \mathcal{A}(\Gamma_{\mathcal{S}})$ . Fix  $\mathcal{C} \in \Gamma_{\mathcal{S}}$  such that  $\mathcal{A}(\mathcal{C}) = \mathbf{Z}$ . Then, we would have  $\mathbf{C}_{k,l}^i = \mathbf{Z}_{k,l}$  for all  $i \in \llbracket r \rrbracket$ ,  $(k,l) \in \text{obs}_i(\mathcal{S})$ . Following the completion algorithm, since the algorithm breaks the loop because of incompatibility, there exists an index  $(k,l)$  such that  $\sum_i \mathbf{C}_{k,l}^i \neq \mathbf{Z}_{k,l}$ . This is a contradiction.

For the second assertion, one verifies that, at initialization, for each index  $(k,l)$ , whenever all entries  $\mathbf{C}_{k,l}^i$ ,  $1 \leq i \leq r$ , are not missing, the sum  $\sum_i \mathbf{C}_{k,l}^i$  is equal to  $\mathbf{Z}_{k,l}$ . This property is preserved at the end of each loop of the algorithm, as by assumption, the algorithm did not break the loop because of incompatibility. Since there is no missing values in the output of the algorithm, we obtain  $\mathcal{A}(\mathcal{C}) = \mathbf{Z}$ . Uniqueness comes simply from the fact that each completed value in the algorithm is unique. ■

In other words, when [Algorithm 2.1](#), run with  $\mathbf{Z}$  and  $\mathcal{S}$  as inputs, returns an output  $\mathcal{C}$  without remaining missing values,  $\mathcal{C}$  is guaranteed to be the P-unique EMD of  $\mathbf{Z}$  in  $\Gamma_{\mathcal{S}}$ . Moreover in this case the PS-unique EMF of  $\mathbf{Z}$  in  $\mathcal{S}$  can be constructed from  $\mathcal{C}$ , where  $\mathcal{S}$  is the unique pair of supports such that  $\varphi(\mathcal{S}) = \mathcal{S}$ . Indeed, one easily constructs  $(\mathbf{X}, \mathbf{Y}) \in \text{IC}_{\mathcal{S}}$  such that  $\mathbf{X}_i \mathbf{Y}_i^\top = \mathbf{C}^i$ . We illustrate an application of this algorithm to the supports of [Figure 1](#) / [Example 2.20](#) in [Figure 4](#).

From [Theorem 2.15](#), we can deduce that one way to guarantee that [Algorithm 2.1](#) returns an output without remaining missing values is to enforce closability on the input tuple of rank-one supports  $\mathcal{S}$ .

**Lemma 2.22.** *Suppose that  $\mathcal{S}$  is closable. Then, for any  $\mathbf{Z} \in \mathcal{A}(\Gamma_{\mathcal{S}}^{(=)})$ , the output of [Algorithm 2.1](#) with inputs  $\mathcal{S}$  and  $\mathbf{Z}$  is an  $r$ -tuple of rank-one contributions without remaining missing values.*

*Proof.* This is in essence a corollary of [Theorem 2.15](#). ■

$$\begin{pmatrix} 0 & 1 & 2 & 0 \\ 1 & 2 & 2 & 0 \\ 2 & 6 & 5 & 6 \\ 3 & 5 & 2 & 4 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ ? & ? & 0 & 0 \\ ? & ? & 0 & 0 \\ ? & ? & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & ? & ? & 0 \\ 0 & ? & ? & 0 \\ 0 & ? & ? & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & ? & ? & ? \\ 0 & ? & ? & ? \end{pmatrix}$$

(a) Initialization.

$$\begin{pmatrix} 0 & 1 & 2 & 0 \\ 1 & 2 & 2 & 0 \\ 2 & 6 & 5 & 6 \\ 3 & 5 & 2 & 4 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & ? & 0 & 0 \\ 2 & ? & 0 & 0 \\ 3 & ? & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 2 & 0 \\ 0 & ? & 2 & 0 \\ 0 & ? & ? & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & ? & ? & 6 \\ 0 & ? & 2 & 4 \end{pmatrix}$$

(b) Filling values in observable supports.

$$\begin{pmatrix} 0 & 1 & 2 & 0 \\ 1 & 2 & 2 & 0 \\ 2 & 6 & 5 & 6 \\ 3 & 5 & 2 & 4 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & ? & 0 & 0 \\ 2 & ? & 0 & 0 \\ 3 & ? & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 2 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & ? & ? & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & ? & 3 & 6 \\ 0 & ? & 2 & 4 \end{pmatrix}$$

(c) Completion operation inside each bipartite graph.

$$\begin{pmatrix} 0 & 1 & 2 & 0 \\ 1 & 2 & 2 & 0 \\ 2 & 6 & 5 & 6 \\ 3 & 5 & 2 & 4 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 2 & ? & 0 & 0 \\ 3 & ? & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 2 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & ? & 2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & ? & 3 & 6 \\ 0 & ? & 2 & 4 \end{pmatrix}$$

(d) Completion operation across bipartite graphs.

$$\begin{pmatrix} 0 & 1 & 2 & 0 \\ 1 & 2 & 2 & 0 \\ 2 & 6 & 5 & 6 \\ 3 & 5 & 2 & 4 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ 3 & 3 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 2 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & ? & 3 & 6 \\ 0 & ? & 2 & 4 \end{pmatrix}$$

(e) Completion operation inside each bipartite graph.

$$\begin{pmatrix} 0 & 1 & 2 & 0 \\ 1 & 2 & 2 & 0 \\ 2 & 6 & 5 & 6 \\ 3 & 5 & 2 & 4 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ 3 & 3 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 2 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 3 & 3 & 6 \\ 0 & 2 & 2 & 4 \end{pmatrix}$$

(f) Completion operation across bipartite graphs and output.

Figure 4: Application of [Algorithm 2.1](#) on the supports of [Figure 1](#) / [Example 2.20](#). We represent each contribution  $\mathcal{C}^i$  at several steps of the completion algorithm. The output does not have remaining values, so by [Proposition 2.21](#), the obtained tuple of rank-one contributions is identifiable in the EMD with fixed support.

**Algorithm 2.1** Fixed-support exact matrix decomposition algorithm**Require:** Observed matrix  $\mathbf{Z} \in \mathbb{C}^{m \times n}$ ,  $r$ -tuple of rank-one matrices  $\mathcal{S} = (\mathcal{S}^i)_{i=1}^r$ **Ensure:** Rank-one contributions  $\mathcal{C}$ , with possible missing values “?”

```

1: For  $1 \leq i \leq r$ , set  $\mathcal{C}^i_{k,l} = \begin{cases} \mathbf{Z}_{k,l} & \text{if } (k,l) \in \text{obs}_i(\mathcal{S}) \\ 0 & \text{if } (k,l) \notin \mathcal{S}^i \\ \text{“?”} & \text{otherwise} \end{cases}$ 
2: while completion of “?” in  $\mathcal{C}$  is still possible do
3:   for  $i = 1$  to  $r$  do /* Completion inside each bipartite graph */
4:     for each submatrix  $\begin{bmatrix} a_1 & a_2 \\ a_4 & a_3 \end{bmatrix}$  in  $\mathcal{C}^i$  do
5:       if there exist three non missing entries and one missing entry  $a_i = \text{“?”}$  then
6:          $i' \leftarrow (i + 2) \bmod 4$ 
7:         if  $a_{i'} \neq 0$  then
8:            $a_i \leftarrow \frac{a_k \times a_j}{a_{i'}}$  with  $\{k, j\}$  the remaining indices in  $\llbracket 4 \rrbracket \setminus \{i, i'\}$ 
9:         end if
10:      end if
11:    end for
12:  end for
13:  for  $(k, l) \in \llbracket m \rrbracket \times \llbracket n \rrbracket$  do /* Checking compatibility of the completion */
14:    if  $\mathcal{C}^i_{k,l} \neq \text{“?”}$  for each  $i \in \llbracket r \rrbracket$  and  $\mathbf{Z}_{k,l} \neq \sum_i \mathcal{C}^i_{k,l}$  then
15:      return break the “while” loop because of incompatibility
16:    end if
17:  end for
18:  for  $(k, l) \in \llbracket m \rrbracket \times \llbracket n \rrbracket$  do /* Completion accross bipartite graphs */
19:    if there exists  $i$  such that  $\mathcal{C}^i_{k,l} = \text{“?”}$  and  $\mathcal{C}^j_{k,l} \neq \text{“?”}$  for all  $j \neq i$  then
20:       $\mathcal{C}^i_{k,l} \leftarrow \mathbf{Z}_{k,l} - \sum_{j \neq i} \mathcal{C}^j_{k,l}$ 
21:    end if
22:  end for
23: end while
24: return Rank-one contributions  $\mathcal{C}$ 

```

We also mention that, [Algorithm 2.1](#) will be used in [section 3](#) in a hierarchical factorization method, to successively recover the multiple sparse factors  $\mathbf{X}_L, \dots, \mathbf{X}_1$  from  $\mathbf{Z} = \mathbf{X}^L \dots \mathbf{X}^1$ , when considering some well-chosen sparsity constraints on the factors.

**2.4. Application for the DFT, DST and DCT matrices.** The DFT matrix, DCT matrix (type II) and DST matrix (type II) of size  $N$  are defined by:

$$(2.17) \quad \mathbf{DFT}_N := (\omega_N^{(k-1)(l-1)})_{k,l \in \llbracket N \rrbracket}, \quad \text{where } \omega_N := e^{-i \frac{2\pi}{N}},$$

$$(2.18) \quad \mathbf{DCT}_N := \left( \cos \left[ \frac{\pi}{N} \left( l - \frac{1}{2} \right) (k - 1) \right] \right)_{k,l \in \llbracket N \rrbracket},$$

$$(2.19) \quad \mathbf{DST}_N := \left( \sin \left[ \frac{\pi}{N} \left( l - \frac{1}{2} \right) k \right] \right)_{k,l \in \llbracket N \rrbracket}.$$

We start by showing that the sufficient conditions given by [Proposition 2.6](#) to identify the rank-one supports constraint among a specific family of sparsity patterns are verified for the exact matrix decomposition of the DFT, DCT and DST matrix.

[Proposition 2.23.](#) *Let  $N = 2^L$  ( $L \geq 2$ ). Using the notations introduced in (2.3), define  $\Omega := \Lambda_{\text{col}}^{N/2} \times \Lambda_{\text{col}}^2$  the family of pairs of left and right supports of size  $N \times N$  which are respectively  $N/2$ -sparse by column and 2-sparse by column. Let  $\mathbf{Z}$  be either  $\mathbf{DFT}_N$ ,  $\mathbf{DCT}_N$ , or  $\mathbf{DST}_N$ , and  $\mathcal{C} \in \Gamma_\Omega$  such that  $\mathbf{Z} = \mathcal{A}(\mathcal{C})$ . Then, the sufficient conditions of [Proposition 2.6](#) on  $\mathcal{C}$  and  $\Omega$  are verified.*

[Remark 2.24.](#) We conjecture that similar results for the DCT and DST of type I and III can be established, but not for type IV. For the sake of conciseness, we only give in this paper identifiability results for the DCT-II and DST-II matrices with their formal proof, as type II is the most common type for DCT and DST. The extension for DCT and DST matrices of type I and III is therefore left as future work.

*Informal proof.* Let us start by giving an informal proof for the case  $\mathbf{Z} = \mathbf{DFT}_4$ , as it illustrates concretely the assumptions of [Proposition 2.6](#). The only partition of  $\text{supp}(\mathbf{DFT}_4)$  into four rank-one supports  $(\mathcal{S}^1, \dots, \mathcal{S}^4)$ , which have at most 2 nonzero rows and 2 nonzero columns, and such that the submatrix of  $\mathbf{DFT}_4$  restricted to  $\mathcal{S}^i$  ( $1 \leq i \leq 4$ ) is of rank one, is the following one:

$$\begin{pmatrix} \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} \\ \mathbf{1} & \mathbf{i} & -\mathbf{1} & -\mathbf{i} \\ \mathbf{1} & -\mathbf{1} & \mathbf{1} & -\mathbf{1} \\ \mathbf{1} & -\mathbf{i} & -\mathbf{1} & \mathbf{i} \end{pmatrix},$$

where the entries of the same rank-one supports  $\mathcal{S}^i$  are represented by the same color. Here,  $\mathcal{S}^1 = \{1, 3\} \times \{1, 3\}$ ,  $\mathcal{S}^2 = \{2, 4\} \times \{1, 3\}$ ,  $\mathcal{S}^3 = \{1, 3\} \times \{2, 4\}$ ,  $\mathcal{S}^4 = \{2, 4\} \times \{2, 4\}$ .  $\blacksquare$

The proof for  $\mathbf{Z} = \mathbf{DFT}_N$  is provided in [10, Chapter 7, Section 7.4]. The proof for  $\mathbf{Z} = \mathbf{DCT}_N$  or  $\mathbf{Z} = \mathbf{DST}_N$  is deferred to [Appendix B](#). The main idea of the proof for  $\mathbf{Z} = \mathbf{DCT}_N$  or  $\mathbf{Z} = \mathbf{DST}_N$  is that for any  $\mathcal{C} \in \Gamma_\Omega$  such that  $\mathbf{DCT}_N = \mathcal{A}(\mathcal{C})$  or  $\mathbf{DST}_N = \mathcal{A}(\mathcal{C})$ , each rank-one support  $\text{supp}(\mathcal{C}^i)$ ,  $1 \leq i \leq N$ , corresponds to a set of  $N/2$  rows with indices of the same parity, and of two columns with indices  $\{l_1^{(i)}, l_2^{(i)}\}$  such that  $(l_1^{(i)} + l_2^{(i)} - 1) = N$ . An illustration of these supports is shown in [Figure 5](#).

We are now able to show that the DFT, DCT and DST admit a unique factorization in the family  $\Omega := \Lambda_{\text{col}}^{N/2} \times \Lambda_{\text{col}}^2$ . Hence, our work generalizes the arguments used in [10, Chapter 7] to prove that the sparse factorization of DFT matrix into two factors is identifiable.

[Theorem 2.25.](#) *Denote  $\mathbf{Z}$  as either the DFT matrix, the DCT matrix (type II), or the DST matrix (type II) of size  $N = 2^L$  ( $L \geq 2$ ). Then,  $\mathbf{Z}$  has a PS-unique EMF in  $\Omega := \Lambda_{\text{col}}^{N/2} \times \Lambda_{\text{col}}^2$ , which is the family of left and right supports of size  $N \times N$  which are respectively  $\frac{N}{2}$ -sparse by column and 2-sparse by column. Moreover,  $\mathbf{DFT}_N$  also admits a PS-unique EMF in  $\Lambda_{\text{col}}^2 \times \Lambda_{\text{col}}^{N/2}$ .*

[Remark 2.26.](#)  $\mathbf{DFT}_N$  also admits a PS-unique EMF in  $\Lambda_{\text{col}}^2 \times \Lambda_{\text{col}}^{N/2}$  because  $\mathbf{DFT}_N^\top = \mathbf{DFT}_N$  by its definition in (2.17).

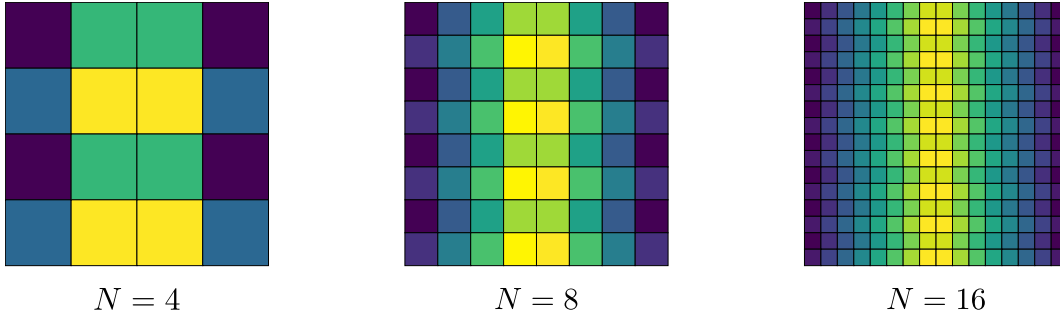


Figure 5: Illustration of the supports  $\{\text{supp}(\mathbf{C}^i)\}_{i=1}^N$ , where  $\mathcal{C}$  is an EMD of  $\mathbf{DCT}_N$  or  $\mathbf{DST}_N$  in  $\Gamma_\Omega$ , with  $\Omega := \Lambda_{\text{col}}^{N/2} \times \Lambda_{\text{col}}^2$ , where  $\Lambda_{\text{col}}^k$  has been defined in (2.3). Each support  $\text{supp}(\mathbf{C}^i)$ ,  $i \in \llbracket N \rrbracket$ , is associated to a distinct color and all indices belonging to this support bear this color.

*Proof.* Suppose there exists  $(\mathbf{X}, \mathbf{Y}) \in \Sigma_\Omega$  such that  $\mathbf{Z} = \mathbf{X}\mathbf{Y}^\top$ . Then, denoting  $\mathcal{C} := \varphi(\mathbf{X}, \mathbf{Y})$ ,  $\mathcal{C} \in \Gamma_\Omega$ , with naturally  $\mathbf{Z} = \mathcal{A}(\mathcal{C})$ . Then, by Proposition 2.23, the sufficient conditions of Proposition 2.6 are satisfied, so the supports  $\{\text{supp}(\mathbf{C}^i)\}_{i=1}^r$  are pairwise disjoint, and are identifiable in the sense of condition (i) in Proposition 2.5. We now show that condition (ii) of Proposition 2.5, *i.e.* fixed-support identifiability, is verified. Let  $\mathcal{S} \in \varphi(\Omega)$  such that  $\mathcal{C} \in \Gamma_{\mathcal{S}}$ . As the cardinal of  $\text{supp}(\mathbf{C}^i)$  is equal to  $N$  for all  $i \in \llbracket r \rrbracket$  (because otherwise,  $\|\mathcal{A}(\mathcal{C})\|_0 < \|\mathbf{Z}\|_0$ ),  $\mathcal{C} \in \Gamma_{\mathcal{S}}^{(=)}$ . A fortiori, the supports  $\{\mathcal{S}^i\}_{i=1}^r$  are pairwise disjoint, so  $\mathcal{S}$  is closable by Definition 2.14. By Theorem 2.15,  $\mathcal{C} \in \mathcal{U}(\Gamma_{\mathcal{S}})$ . Hence, by Proposition 2.5,  $\mathcal{C} \in \mathcal{U}(\Gamma_\Omega)$ . It remains to show  $(\mathbf{X}, \mathbf{Y}) \in \text{IC}_\Omega \cap \text{MC}_\Omega$ . If  $(\mathbf{X}, \mathbf{Y}) \notin \text{IC}_\Omega$ , there would exist  $i$  such that  $\mathbf{C}^i = 0$ , which is a contradiction. We now show  $(\mathbf{X}, \mathbf{Y}) \in \text{MC}_\Omega$ . Let  $\mathcal{S} \in \Omega$  such that  $(\mathbf{X}, \mathbf{Y}) \in \Sigma_{\mathcal{S}}$ , and denote  $\mathcal{S} := \varphi(\mathcal{S})$ . Then,  $\mathcal{C} \in \Gamma_{\mathcal{S}}$ , but we previously showed that this implies  $\mathcal{C} \in \Gamma_{\mathcal{S}}^{(=)}$ . This yields  $\text{colsupp}(\mathbf{X}) = \text{colsupp}(\mathbf{S}^L)$  and  $\text{colsupp}(\mathbf{Y}) = \text{colsupp}(\mathbf{S}^R)$ . By Theorem 2.4, we conclude that  $(\mathbf{X}, \mathbf{Y}) \in \mathcal{U}(\Sigma_\Omega)$ . ■

We also show a negative result of identifiability for the sparse matrix factorization into two factors of the Hadamard matrix in  $\Sigma_\Omega$ , with  $\Omega := \Lambda_{\text{col}}^{N/2} \times \Lambda_{\text{col}}^2$ , which is reminiscent of some negative examples showing that the sparse factors of the Hadamard matrix cannot be uniquely recovered with a hierarchical factorization method, given the sparsity constraint  $\Omega$  [8].

**Proposition 2.27.** *Define the Hadamard matrix  $\mathbf{H}_{2^L}$  ( $L \geq 2$ ) as:*

$$(2.20) \quad \mathbf{H}_{2^L} := \begin{pmatrix} \mathbf{H}_{2^{L-1}} & \mathbf{H}_{2^{L-1}} \\ \mathbf{H}_{2^{L-1}} & -\mathbf{H}_{2^{L-1}} \end{pmatrix} \in \mathbb{B}^{2^L \times 2^L}$$

where  $\mathbf{H}_1 := (1)$ . Then, the Hadamard matrix  $\mathbf{H}_N$  with  $N = 2^L$  for  $L \geq 2$  does not admit a unique sparse matrix factorization in  $\Omega := \Lambda_{\text{col}}^{N/2} \times \Lambda_{\text{col}}^2$ .

*Proof.* This is simply because for each  $\mathcal{C}' \in \Gamma_\Omega$  such that  $\mathcal{A}(\mathcal{C}') = \mathbf{Z}$ , the supports  $\{\text{supp}(\mathbf{C}'^i)\}_{i=1}^r$  are pairwise disjoint (see Lemma B.1), and there exists several distinct partitions  $\{\mathcal{S}^i\}_{i=1}^r$  of  $\text{supp}(\mathbf{H}_N)$  with  $\mathcal{S} \in \varphi^{\text{cp}}(\Omega)$  such that for all  $i \in \llbracket r \rrbracket$ ,  $(\mathbf{Z}_{k,l})_{(k,l) \in \mathcal{S}^i}$  is of rank at



most one. Therefore, we can easily find  $\mathcal{C}, \mathcal{C}' \in \Gamma_\Omega$  such that  $\mathcal{C} \not\sim \mathcal{C}'$ , but  $\mathcal{A}(\mathcal{C}) = \mathbf{H}_\mathbf{N} = \mathcal{A}(\mathcal{C}')$ . This means that  $\mathbf{H}_\mathbf{N}$  does not admit a P-unique EMD in  $\Gamma_\Omega$ . By [Theorem 2.4](#), we conclude that  $\mathbf{H}_\mathbf{N}$  does not admit a PS-unique EMF in  $\Sigma_\Omega$ .  $\blacksquare$

**3. Identifiability in multilayer sparse matrix factorization.** The analysis with two-factors presented in the companion paper [\[14\]](#), and illustrated in the previous section for the DFT, DCT and DST matrices, can provide tools for analyzing identifiability in exact sparse matrix factorization with  $L > 2$  factors. This extension is based on a hierarchical factorization method [\[11, 10\]](#) that will be detailed below. In this section, we establish some identifiability results in the multilayer case in a specific setting where the enforced sparsity constraints on the factors to recover are the so-called *butterfly supports* (see [Definition 3.4](#) below). Let us highlight here that, in contrary to our analysis with two factors where sparsity constraints on the factors were encoded by a *family* of sparsity patterns, the study in the multilayer case proposed in this paper is limited to the case where this family is reduced to a *singleton*. In other words, we study identifiability in the multilayer case with *fixed-support* sparsity constraint. The possible extension of the proposed analysis to a *family* of sparsity patterns with its difficulties will be discussed at the end of the section.

The butterfly supports are defined from the supports of the butterfly factors of the DFT matrix [\[12, 5\]](#). Let us introduce the *multilayer exact matrix factorization* (MEMF) problem, which generalizes the problem (EMF) with two factors. Given a number of layers  $L > 2$ , a matrix  $\mathbf{Z} \in \mathbb{C}^{N_L \times N_0}$ , and a feasible set of  $L$ -tuples of factors  $\Sigma \subseteq \mathbb{C}^{N_L \times N_{L-1}} \times \dots \times \mathbb{C}^{N_1 \times N_0}$  for some integers  $N_1, \dots, N_{L-1}$ , the MEMF problem for  $\mathbf{Z}$  in  $\Sigma$  is the following multilinear inverse problem:

$$\text{(MEMF)} \quad \text{find, if possible, } (\mathbf{X}_L, \dots, \mathbf{X}_1) \in \Sigma \text{ such that } \mathbf{Z} = \mathbf{X}_L \mathbf{X}_{L-1} \dots \mathbf{X}_1.$$

We only provide an analysis of identifiability in the MEMF for a feasible set  $\Sigma$  encoding *fixed* supports constraints on the factors. For any  $L$ -tuple of supports  $\mathbf{S} := (\mathbf{S}^L, \dots, \mathbf{S}^1)$ , define:

$$(3.1) \quad \Sigma_{\mathbf{S}} := \Sigma_{\mathbf{S}^L} \times \dots \times \Sigma_{\mathbf{S}^1}.$$

The notion of PS-uniqueness of an EMF (with two factors) has been introduced in order to take into account both of permutation and scaling ambiguities. It is therefore possible to generalize this definition to the multilayer case. However, when fixing the butterfly supports as the sparsity constraints, we can show a stronger property than PS-uniqueness, which is identifiability up to scaling ambiguities only. This is due to the fact that the structure of the butterfly supports removes any kind of permutation ambiguities at each factor, as it will be shown in the next subsections.

**Definition 3.1 (S-uniqueness in MEMF).** *For any feasible set  $\Sigma$  of  $L$ -tuples of factors, the tuple  $(\mathbf{X}_L, \dots, \mathbf{X}_1)$  is the  $\mathbf{S}$ -unique MEMF of  $\mathbf{Z} := \mathbf{X}_L \mathbf{X}_{L-1} \dots \mathbf{X}_1$  in  $\Sigma$ , if any solution  $(\mathbf{X}'_L, \dots, \mathbf{X}'_1)$  to (MEMF) of  $\mathbf{Z}$  in  $\Sigma$  is equivalent to  $(\mathbf{X}_L, \dots, \mathbf{X}_1)$ , written  $(\mathbf{X}_L, \dots, \mathbf{X}_1) \sim_s (\mathbf{X}'_L, \dots, \mathbf{X}'_1)$ , in the sense that there exist diagonal matrices with nonzero diagonal entries  $\mathbf{D}_1 \in \mathcal{D}_{N_1}$ ,  $\mathbf{D}_2 \in \mathcal{D}_{N_2}$ ,  $\dots$ ,  $\mathbf{D}_{L-1} \in \mathcal{D}_{N_{L-1}}$  such that for all  $\ell \in \llbracket L \rrbracket$ , we have  $\mathbf{X}'_\ell = \mathbf{D}_\ell^{-1} \mathbf{X}_\ell \mathbf{D}_{\ell-1}$ , with the convention  $\mathbf{D}_L = \mathbf{I}_{N_L}$  and  $\mathbf{D}_0 = \mathbf{I}_{N_0}$ .*

*Remark 3.2.* The formal definition of PS-uniqueness of an MEMF is analogue to the one above, with permutation and scaling equivalence relation  $\sim$  instead of  $\sim_s$ , and generalized permutations instead of diagonal matrices with nonzero diagonal entries.

In the following, when dealing with block diagonal matrices, we write “0” to specify that other entries than the ones in the blocks are zero.

**3.1. Butterfly factorization of the DFT matrix.** Recall that the DFT matrix of size  $N \times N$  with  $N = 2^L$  is denoted as  $\mathbf{DFT}_N$ , and has been introduced in (2.17). The butterfly factorization of  $\mathbf{DFT}_N$  [5] relies on the recursive relation:

$$(3.2) \quad \mathbf{DFT}_N = \mathbf{B}_N \begin{pmatrix} \mathbf{DFT}_{N/2} & 0 \\ 0 & \mathbf{DFT}_{N/2} \end{pmatrix} \mathbf{P}_N,$$

where  $\mathbf{P}_N \in \mathbb{B}^{N \times N}$  is the permutation matrix which sorts the odd indices, then the even indices (e.g., for  $N = 4$ , it permutes  $[1, 2, 3, 4]$  to  $[1, 3, 2, 4]$ ), and

$$(3.3) \quad \mathbf{B}_N := \begin{pmatrix} \mathbf{I}_{N/2} & \mathbf{A}_{N/2} \\ \mathbf{I}_{N/2} & -\mathbf{A}_{N/2} \end{pmatrix},$$

with  $\mathbf{A}_{N/2}$  the diagonal matrix of size  $\frac{N}{2} \times \frac{N}{2}$  with diagonal entries  $1, \omega_N, \omega_N^2, \dots, \omega_N^{\frac{N}{2}-1}$ . Applying recursively (3.2) to the block  $\mathbf{DFT}_{N/2}$ , we obtain the butterfly factorization of  $\mathbf{DFT}_N$  written as:

$$(3.4) \quad \mathbf{DFT}_N = \mathbf{B}_N \begin{pmatrix} \mathbf{B}_{N/2} & 0 \\ 0 & \mathbf{B}_{N/2} \end{pmatrix} \cdots \begin{pmatrix} \mathbf{B}_4 & & 0 \\ & \ddots & \\ 0 & & \mathbf{B}_4 \end{pmatrix} \begin{pmatrix} \mathbf{B}_2 & & & \\ & \mathbf{B}_2 & & 0 \\ & & \ddots & \\ 0 & & & \mathbf{B}_2 \\ & & & & \mathbf{B}_2 \end{pmatrix} \mathbf{R}_N$$

where  $\mathbf{R}_N \in \mathbb{B}^{N \times N}$  is the so-called *bit-reversal permutation matrix*, defined by:

$$(3.5) \quad \mathbf{R}_N = \mathbf{Q}_1 \mathbf{Q}_2 \cdots \mathbf{Q}_L, \quad \text{where } \mathbf{Q}_\ell := \mathbf{I}_{N/2^\ell} \otimes \mathbf{P}_{2^\ell} = \begin{pmatrix} \mathbf{P}_{2^\ell} & & 0 \\ & \ddots & \\ 0 & & \mathbf{P}_{2^\ell} \end{pmatrix}, \quad \forall 1 \leq \ell \leq L.$$

**Definition 3.3 (Butterfly factors of the DFT matrix).** We define the butterfly factors  $(\mathbf{F}_\ell)_{\ell=1}^L$  of size  $N = 2^L$  of the DFT matrix  $\mathbf{DFT}_N$  as the factors:

$$\mathbf{F}_L := \mathbf{B}_N, \quad \mathbf{F}_{L-1} := \begin{pmatrix} \mathbf{B}_{N/2} & 0 \\ 0 & \mathbf{B}_{N/2} \end{pmatrix}, \quad \dots, \quad \mathbf{F}_2 := \begin{pmatrix} \mathbf{B}_4 & & 0 \\ & \ddots & \\ 0 & & \mathbf{B}_4 \end{pmatrix},$$

$$\mathbf{F}_1 := \begin{pmatrix} \mathbf{B}_2 & & & \\ & \mathbf{B}_2 & & 0 \\ & & \ddots & \\ 0 & & & \mathbf{B}_2 \\ & & & & \mathbf{B}_2 \end{pmatrix},$$

in such a way that  $\mathbf{DFT}_N = \mathbf{F}_L \mathbf{F}_{L-1} \cdots \mathbf{F}_1 \mathbf{R}_N$ .

**Definition 3.4 (Butterfly supports).** *The butterfly supports of size  $N = 2^L$  are defined as the  $L$ -tuple of supports  $\mathbf{S} := (\mathbf{S}^L, \dots, \mathbf{S}^1)$  where  $\mathbf{S}^\ell := \text{supp}(\mathbf{F}_\ell)$  for  $1 \leq \ell \leq L$ , and  $(\mathbf{F}_L, \dots, \mathbf{F}_1)$  are the butterfly factors.*

**3.2. Some properties of the butterfly supports.** Figure 6 is an illustration of the butterfly supports. They are 2-regular [9], *i.e.* they have at most 2 nonzero entries per row and per column, but they are also block diagonal (except the leftmost butterfly support which has a single block that covers the matrix entirely).

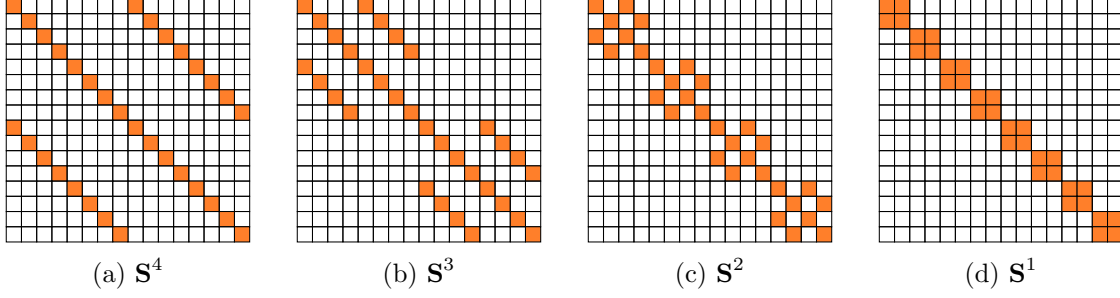


Figure 6: Butterfly supports of size  $N = 16$ . Nonzero entries are in color, and zero entries are in white.

Consequently, given  $(\mathbf{X}_L, \dots, \mathbf{X}_1) \in \Sigma_{\mathbf{S}}$  where  $\mathbf{S}$  is the tuple of butterfly supports, the partial product of any consecutive factors  $\mathbf{X}_q \dots \mathbf{X}_p$  ( $1 \leq p \leq q \leq L$ ) has a very precise structure as detailed in the following lemma. Figure 7 illustrates this structure on some examples of partial products for the case of the butterfly supports of size  $N = 16$ .

**Lemma 3.5.** *Let  $\mathbf{S} := (\mathbf{S}^L, \dots, \mathbf{S}^1)$  be the butterfly supports of size  $N = 2^L$ . Then, for any tuple  $(\mathbf{X}_L, \dots, \mathbf{X}_1) \in \Sigma_{\mathbf{S}}$ , we have  $\text{supp}(\mathbf{X}_q \dots \mathbf{X}_p) \subseteq \mathbf{W}^{[q;p]}$  for any  $1 \leq p \leq q \leq L$ , where:*

$$(3.6) \quad \mathbf{W}^{[q;p]} := \mathbf{I}_{N/2^q} \otimes \mathbf{V}^{q,p} = \begin{pmatrix} \mathbf{V}^{q,p} & & 0 \\ & \ddots & \\ 0 & & \mathbf{V}^{q,p} \end{pmatrix} \in \mathbb{B}^{N \times N}, \text{ and}$$

$$(3.7) \quad \mathbf{V}^{q,p} := \mathbf{U}_{2^{q-p+1}} \otimes \mathbf{I}_{2^{p-1}} = \begin{pmatrix} \mathbf{I}_{2^{p-1}} & \dots & \mathbf{I}_{2^{p-1}} \\ \vdots & & \vdots \\ \mathbf{I}_{2^{p-1}} & \dots & \mathbf{I}_{2^{p-1}} \end{pmatrix} \in \mathbb{B}^{2^q \times 2^q},$$

denoting  $\mathbf{U}_n \in \mathbb{B}^{n \times n}$  as the binary matrix full of ones for any  $n$ . Moreover, if we suppose in addition  $\text{supp}(\mathbf{X}_\ell) = \mathbf{S}^\ell$  for all  $1 \leq \ell \leq L$ , then  $\text{supp}(\mathbf{X}_q \dots \mathbf{X}_p) = \mathbf{W}^{[q;p]}$  for any  $1 \leq p \leq q \leq L$ .

*Proof.* We start by the case where  $q = L$ . Let us show by backward induction that  $\text{supp}(\mathbf{X}_L \dots \mathbf{X}_\ell) \subseteq \mathbf{V}^{L,\ell}$  for any  $1 \leq \ell \leq L$ . This is true for  $\ell = L$ , because by definition,  $\text{supp}(\mathbf{X}_L) \subseteq \mathbf{S}^L = \mathbf{V}^{L,L}$ . Let  $2 \leq \ell \leq L$ , and suppose that  $\text{supp}(\mathbf{X}_L \dots \mathbf{X}_\ell) \subseteq \mathbf{V}^{L,\ell}$ , which is the matrix full of blocks  $\mathbf{I}_{2^{\ell-1}}$ . Since  $\mathbf{S}^{\ell-1}$  is block diagonal with blocks of size  $2^{\ell-1} \times 2^{\ell-1}$

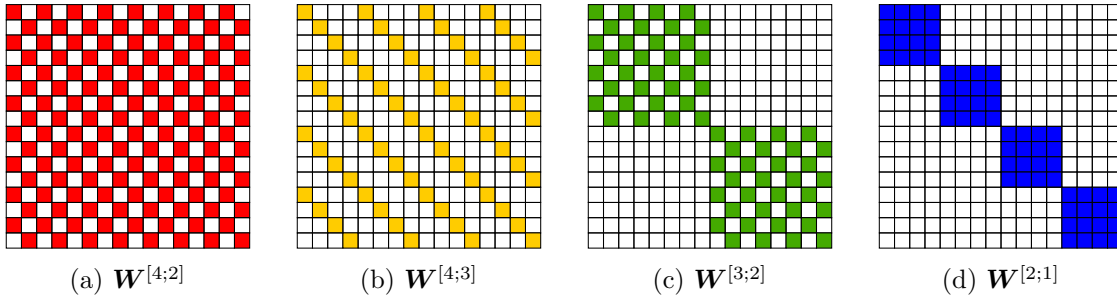


Figure 7: Examples of supports  $\mathbf{W}^{[q;p]}$  ( $1 \leq p \leq q \leq 4$ ) of size  $N \times N$  with  $N = 16$ . Nonzero entries are in color, and zero entries are in white.

equal to  $\mathbf{U}_2 \otimes \mathbf{I}_{2^{\ell-2}} = \begin{pmatrix} \mathbf{I}_{2^{\ell-2}} & \mathbf{I}_{2^{\ell-2}} \\ \mathbf{I}_{2^{\ell-2}} & \mathbf{I}_{2^{\ell-2}} \end{pmatrix}$ , the product  $\mathbf{V}^{L,\ell} \mathbf{S}^{\ell-1}$  is equal to  $\mathbf{V}^{L,\ell-1}$ . Consequently, because of the block diagonal structure of  $\mathbf{S}^{\ell-1}$ , we have  $\text{supp}(\mathbf{X}_L \dots \mathbf{X}_\ell \mathbf{X}_{\ell-1}) \subseteq \mathbf{V}^{L,\ell} \mathbf{S}^{\ell-1} = \mathbf{V}^{L,\ell-1}$ . This ends the proof by induction.

Now, in the case where  $q < L$ , the butterfly support  $\mathbf{S}^q$  of size  $N \times N$  is block diagonal, where each block is simply the leftmost butterfly support of size  $2^q \times 2^q$ . Applying the previous case on each of these blocks of size  $2^q \times 2^q$  yields the claimed result, *i.e.*  $\text{supp}(\mathbf{X}_q \dots \mathbf{X}_p) \subseteq \mathbf{W}^{[q;p]}$  for any  $1 \leq p \leq q \leq L$ . In the case where  $\text{supp}(\mathbf{X}_\ell) = \mathbf{S}^\ell$  for all  $1 \leq \ell \leq L$ , we repeat the proof, by first starting to prove inductively that  $\text{supp}(\mathbf{X}_L \dots \mathbf{X}_\ell) = \mathbf{V}^{L,\ell}$  for any  $1 \leq \ell \leq L$ , and then generalizing the result to  $q < L$ .  $\blacksquare$

This sparsity structure of partial products will be exploited to establish identifiability for the EMF of  $\mathbf{Z}^{[q;p]} := (\mathbf{X}_q \dots \mathbf{X}_{\ell+1})(\mathbf{X}_\ell \dots \mathbf{X}_p)$  into two factors, when fixing the supports  $\mathbf{W}^{[q;\ell+1]}$  and  $\mathbf{W}^{[\ell;p]}$  as the respective sparsity constraint for the left and right factor ( $1 \leq p \leq \ell < q \leq L$ ).

**Lemma 3.6.** *The  $N$ -tuple of rank-one supports  $\varphi(\mathbf{W}^{[q;\ell+1]}, \mathbf{W}^{[\ell;p]})$  has disjoint rank-one supports, for any  $1 \leq p \leq \ell < q \leq L$ . Consequently:*

$$\mathcal{U}_s \left( \Sigma_{(\mathbf{W}^{[q;\ell+1]}, \mathbf{W}^{[\ell;p]})} \right) = \text{IC}_{(\mathbf{W}^{[q;\ell+1]}, \mathbf{W}^{[\ell;p]})} \cap \text{MC}_{(\mathbf{W}^{[q;\ell+1]}, \mathbf{W}^{[\ell;p]})}.$$

*Remark 3.7.* Since the matrix  $\mathbf{W}^{[\ell;p]}$  is symmetric, we write  $\mathbf{W}^{[\ell;p]}$  instead of  $(\mathbf{W}^{[\ell;p]})^\top$ .

*Proof.* Denote  $\mathcal{S} := \varphi(\mathbf{W}^{[q;\ell+1]}, \mathbf{W}^{[\ell;p]})$ . We also write  $I_k := \{(k-1)2^\ell + 1, \dots, k2^\ell - 1\}$  for  $k \in \llbracket \frac{N}{2^\ell} \rrbracket$ . On the one hand, the right support  $\mathbf{W}^{[\ell;p]}$  is block diagonal, with blocks  $\mathbf{V}^{\ell,p}$  of size  $2^\ell \times 2^\ell$ . This means that for  $i \in I_k$  and  $j \in I_{k'}$  with  $k, k' \in \llbracket \frac{N}{2^\ell} \rrbracket$ ,  $k \neq k'$ , the rank-one supports  $\mathcal{S}^i$  and  $\mathcal{S}^j$  are disjoint. On the other hand, viewing column supports as subsets of indices, the columns of the left support  $\mathbf{W}^{[q;\ell+1]}$  restricted to  $I_k$  are pairwise disjoint, by definition of  $\mathbf{W}^{[q;\ell+1]}$ . This means that for  $i, j \in I_k$  ( $k \in \llbracket \frac{N}{2^\ell} \rrbracket$ ), the rank-one supports  $\mathcal{S}^i$  and  $\mathcal{S}^j$  are also disjoint, when  $i \neq j$ . Hence, we obtain the claimed result by applying [Corollary 2.19](#).  $\blacksquare$

**3.3. Hierarchical factorization method.** The hierarchical matrix factorization [11] is a method to recover the sparse factors  $\mathbf{X}_L, \dots, \mathbf{X}_1$  from the product  $\mathbf{Z} = \mathbf{X}_L \dots \mathbf{X}_1$ , by successively recovering the factors through several hierarchical levels. At the first level, the factors  $\mathbf{X}_L$  and  $\mathbf{H}_{L-1} := \mathbf{X}_{L-1} \dots \mathbf{X}_1$  are recovered from  $\mathbf{Z}$ . At the second level, the factors  $\mathbf{X}_{L-1}$  and  $\mathbf{H}_{L-2} := \mathbf{X}_{L-2} \dots \mathbf{X}_1$  are recovered from  $\mathbf{H}_{L-1}$ . The process is repeated recursively, until all the sparse factors have been recovered. Note that at each hierarchical factorization, adequate support constraints are enforced on the factors: they should correspond to sparsity patterns obtained from the partial product of several sparse factors.

When the sparse factorization at each hierarchical level is unique up to scaling ambiguities for these adequate support constraints, the recovered factors  $\mathbf{X}_L, \dots, \mathbf{X}_1$  are shown to be the S-unique MEMF of  $\mathbf{Z}$  for the considered constraints. In our case, we enforce the butterfly support structure on the factors. Based on the previous lemmas and on this hierarchical factorization method, we obtain the main result of this section.

**Theorem 3.8.** *Let  $\mathbf{S} := (\mathbf{S}^L, \dots, \mathbf{S}^1)$  be the butterfly supports of size  $N = 2^L$ . Then, for any  $(\mathbf{X}_L, \mathbf{X}_{L-1}, \dots, \mathbf{X}_1) \in \Sigma_{\mathbf{S}}$  such that  $\text{supp}(\mathbf{X}_\ell) = \mathbf{S}^\ell$  for all  $1 \leq \ell \leq L$ , the factors  $\mathbf{X}_L, \dots, \mathbf{X}_1$  are the S-unique MEMF of  $\mathbf{Z} := \mathbf{X}_L \dots \mathbf{X}_1$  in  $\Sigma_{\mathbf{S}}$ . Moreover, these factors can be recovered from  $\mathbf{Z}$ , up to scaling ambiguities, through a hierarchical factorization method detailed in [Algorithm 3.1](#), where  $\mathbf{Z}$  and any partitioning binary tree  $\mathcal{T}$  of  $\{1, \dots, L\}$  (see [Definition 3.10](#) below) are given as the inputs of the algorithm.*

This theorem generalizes a result suggested in [10, Chapter 7] but not proven formally.

**Corollary 3.9.** *The factors  $(\mathbf{F}_L, \mathbf{F}_{L-1}, \dots, \mathbf{F}_1 \mathbf{R}_N)$  defined in [Definition 3.3](#) are the S-unique MEMF of the DFT matrix of size  $N = 2^L$  in  $\Sigma_{(\mathbf{S}^L, \dots, \mathbf{S}^2, \mathbf{S}^1 \mathbf{R}_N)}$ , where the L-tuple of butterfly supports of size  $N$  is denoted by  $\mathbf{S} := (\mathbf{S}^L, \dots, \mathbf{S}^2, \mathbf{S}^1)$ .*

Before stating [Algorithm 3.1](#), let us mention that, in the specific case of the butterfly support constraint, the hierarchical factorization method previously described can be generalized to any kind of binary tree structure. For instance, in the case of four factors  $\mathbf{X}_4, \dots, \mathbf{X}_1$  of size  $16 \times 16$ , one can perform hierarchical factorization in an alternative fashion, by factorizing  $\mathbf{Z} := \mathbf{X}_4 \mathbf{X}_3 \mathbf{X}_2 \mathbf{X}_1$  into  $\mathbf{X}_4 \mathbf{X}_3$  and  $\mathbf{X}_2 \mathbf{X}_1$  at the first layer, then  $\mathbf{X}_4 \mathbf{X}_3$  into  $\mathbf{X}_4$  and  $\mathbf{X}_3$ , and finally  $\mathbf{X}_2 \mathbf{X}_1$  into  $\mathbf{X}_2$  and  $\mathbf{X}_1$ . Let us give a formal definition for such a binary tree structure that defines the factorization order in the hierarchical factorization method.

**Definition 3.10 (Partitioning binary tree).** *A partitioning binary tree of a set of indices  $\{1, \dots, L\}$  is a binary tree, where nodes are non empty subsets of  $\{1, \dots, L\}$ , which satisfies the following axioms:*

- (i) *Each node is a subset of consecutive indices.*
- (ii) *The root is the set  $\{1, \dots, L\}$ .*
- (iii) *A node is a singleton if, and only if, it is a leaf.*
- (iv) *For a node which is not a leaf, the left and right children form a partition of their parent, in a such way that the indices of the left child are larger than those in the right child.*

**Remark 3.11.** This notion is somehow reminiscent of *dimension partition tree* [6] used to define tree tensor networks.

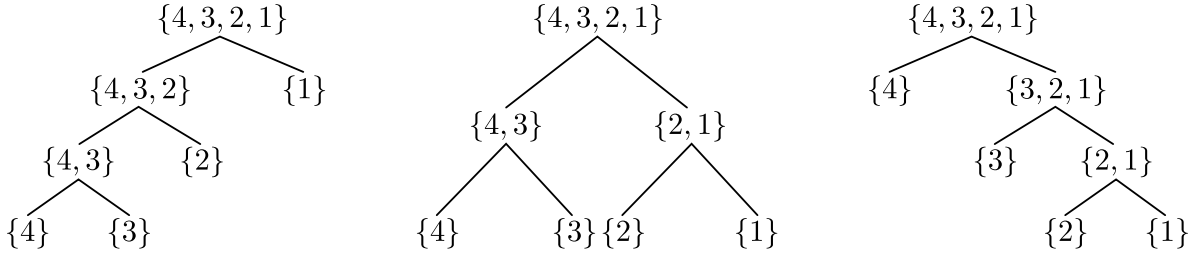


Figure 8: Three possible partitioning binary trees of  $\{1, 2, 3, 4\}$ . Any of these trees given as input of [Algorithm 3.1](#) yields the exact recovery of the sparse factors  $\mathbf{X}_4, \dots, \mathbf{X}_1$  in  $\mathbf{Z} := \mathbf{X}_4 \dots \mathbf{X}_1$ , when  $\text{supp}(\mathbf{X}_\ell)$  are exactly the butterfly supports. Let us highlight here that this result is true for any number of layers  $L$ , and is not limited to the case  $L = 4$ .

Examples of partitioning binary trees are illustrated in [Figure 8](#). We are now able to describe [Algorithm 3.1](#), and give a proof for [Theorem 3.8](#). The key of the proof is the  $S$ -uniqueness of the factorization at each hierarchical level, up to rescaling of columns and rows of the factors: more precisely, at line 8 of [Algorithm 3.1](#),  $(\mathbf{H}_2, \mathbf{H}_1) \in \mathcal{U}_s \left( \Sigma_{(\mathbf{W}^{[q;\ell+1]}, \mathbf{W}^{[\ell;p]})} \right)$  by [Lemma 3.6](#). Formally, we rely on the following lemma.

---

**Algorithm 3.1** Hierarchical factorization method for exact sparse matrix factorization, constrained by the butterfly supports  $(\mathbf{S}^L, \dots, \mathbf{S}^1)$  of size  $N = 2^L$ .

---

**Require:** Matrix  $\mathbf{Z} := \mathbf{X}_q \dots \mathbf{X}_p$  with  $\text{supp}(\mathbf{X}_\ell) = \mathbf{S}^\ell$  for all  $1 \leq p \leq \ell \leq q \leq L$ , and a partitioning binary tree  $\mathcal{T}$  of  $\{p, \dots, q\}$ .

**Ensure:** factors  $\mathbf{X}'_q, \dots, \mathbf{X}'_p$  such that  $(\mathbf{X}'_q, \dots, \mathbf{X}'_p) \sim_s (\mathbf{X}_q, \dots, \mathbf{X}_p)$

- 1: **if** root of  $\mathcal{T}$  is a leaf **then**
  - 2:     **return**  $\mathbf{Z}$  /\* Case where  $q = p$ ,  $\mathbf{Z} = \mathbf{X}_p = \mathbf{X}_q$  \*/
  - 3: **end if**
  - 4:  $(\mathcal{T}_2, \mathcal{T}_1) \leftarrow$  left and right subtrees of the root of  $\mathcal{T}$  /\* Case where  $q > p$  \*/
  - 5:  $\ell \leftarrow$  maximum value in the root of  $\mathcal{T}_1$
  - 6:  $\mathcal{S} \leftarrow \varphi(\mathbf{W}^{[q;\ell+1]}, \mathbf{W}^{[\ell;p]})$
  - 7: **for**  $i = 1$  **to**  $N$  **do**
  - 8:      $\mathbf{C}^i \leftarrow (\mathbf{Z}_{j,k})_{(j,k) \in \mathcal{S}^i}$ , where the support  $\mathcal{S}^i$  is viewed as a subset of indices
  - 9: **end for**
  - 10: set  $(\mathbf{H}_2, \mathbf{H}_1)$  as a pair of factors in  $\varphi^{-1}(\{\mathcal{C}\}) \cap \text{IC}_{(\mathbf{W}^{[q;\ell+1]}, \mathbf{W}^{[\ell;p]})}$
  - 11:  $(\mathbf{X}'_q, \dots, \mathbf{X}'_{\ell+1}) \leftarrow$  result of [Algorithm 3.1](#) with inputs  $\mathbf{H}_2$  and  $\mathcal{T}_2$
  - 12:  $(\mathbf{X}'_\ell, \dots, \mathbf{X}'_p) \leftarrow$  result of [Algorithm 3.1](#) with inputs  $\mathbf{H}_1$  and  $\mathcal{T}_1$
  - 13: **return**  $(\mathbf{X}'_q, \dots, \mathbf{X}'_{\ell+1}, \mathbf{X}'_\ell, \dots, \mathbf{X}'_p)$
- 

**Lemma 3.12.** Let  $\mathbf{S} := (\mathbf{S}^L, \dots, \mathbf{S}^1)$  be the butterfly supports of size  $N = 2^L$ , and let  $(\mathbf{X}_L, \dots, \mathbf{X}_1) \in \Sigma_{\mathbf{S}}$  with  $\text{supp}(\mathbf{X}_\ell) = \mathbf{S}^\ell$ ,  $1 \leq \ell \leq L$ . Then, for any  $1 \leq p \leq \ell < q \leq L$ , for any diagonal matrices  $\mathbf{D}_q, \mathbf{D}_p \in \mathcal{D}_r$ , the matrix  $\mathbf{D}_q^{-1} \mathbf{X}_q \dots \mathbf{X}_{\ell+1} \mathbf{X}_\ell \dots \mathbf{X}_p \mathbf{D}_p$  admits an  $S$ -unique EMF in  $\Sigma_{(\mathbf{W}^{[q;\ell+1]}, \mathbf{W}^{[\ell;p]})}$ , which is  $([\mathbf{D}_q^{-1} \mathbf{X}_q \dots \mathbf{X}_{\ell+1}], [\mathbf{X}_\ell \dots \mathbf{X}_p \mathbf{D}_p]^\top)$

*Proof.* Denote  $\mathbf{H}_2 := \mathbf{X}_q \dots \mathbf{X}_{\ell+1}$ , and  $\mathbf{H}_1^\top := \mathbf{X}_\ell \dots \mathbf{X}_p$ . By Lemma 3.5,  $\text{supp}(\mathbf{H}_2) = \mathbf{W}^{[q;\ell+1]}$ , and  $\text{supp}(\mathbf{H}_1) = \mathbf{W}^{[\ell;p]}$ . The multiplication of a matrix by  $\mathbf{D}_q$  or  $\mathbf{D}_p$  does not change its support, so  $\text{supp}(\mathbf{D}_q^{-1}\mathbf{H}_2) = \text{supp}(\mathbf{H}_2)$  and  $\text{supp}(\mathbf{H}_1^\top \mathbf{D}_p) = \text{supp}(\mathbf{H}_1^\top)$ . In particular,  $\text{colsupp}(\mathbf{D}_q^{-1}\mathbf{H}_2) = \text{colsupp}(\mathbf{D}_p\mathbf{H}_1) = \llbracket N \rrbracket$ , so by Lemma 3.6, we obtain:

$$(\mathbf{D}_q^{-1}\mathbf{H}_2, \mathbf{D}_p\mathbf{H}_1) \in \text{IC}_{(\mathbf{W}^{[q;\ell+1]}, \mathbf{W}^{[\ell;p]})} \cap \text{MC}_{(\mathbf{W}^{[q;\ell+1]}, \mathbf{W}^{[\ell;p]})} = \mathcal{U}_s(\Sigma_{(\mathbf{W}^{[q;\ell+1]}, \mathbf{W}^{[\ell;p]})}). \quad \blacksquare$$

*Proof of Theorem 3.8.* We introduce the following notations. For each node  $n$  in the partitioning binary tree  $\mathcal{T}$ , we write  $q_n$  and  $p_n$  respectively the maximum and minimum index of the corresponding subset of node  $n$ . When  $n$  has children, we write  $\ell_n$  the ‘‘splitting’’ index of node  $n$ , which corresponds to the maximum value of the subset corresponding to its right child.

Let  $(\mathbf{X}_L, \dots, \mathbf{X}_1), (\mathbf{X}'_L, \dots, \mathbf{X}'_1) \in \Sigma_S$  such that  $\mathbf{X}'_L \dots \mathbf{X}'_1 = \mathbf{Z} := \mathbf{X}_L \dots \mathbf{X}_1$ , and suppose that  $(\text{supp}(\mathbf{X}_\ell))_{\ell=1}^L$  are exactly the butterfly supports. For any  $1 \leq p < q \leq L$ , denote  $\mathbf{H}^{[q;p]} = \mathbf{X}_q \dots \mathbf{X}_p$ , and  $\mathbf{H}'^{[q;p]} = \mathbf{X}'_q \dots \mathbf{X}'_p$ . Consider the sequence of nodes  $(n_1, n_2, \dots, n_w)$  in  $\mathcal{T}$  that are obtained after a breadth-first search of the binary tree, where  $w$  is the total number of nodes in  $\mathcal{T}$ . Then, consider the subsequence of  $(n_1, n_2, \dots, n_w)$  where we keep only nodes that are not leaves, and denote them  $n_{\gamma(1)}, \dots, n_{\gamma(v)}$ , with  $v$  the total number of nodes in  $\mathcal{T}$  that are *not* leaves. To simplify the notations, we denote, for  $1 \leq u \leq v$ ,  $\ell_{\gamma(u)} := \ell_{n_{\gamma(u)}}$ ,  $q_{\gamma(u)} := q_{n_{\gamma(u)}}$ , and  $p_{\gamma(u)} := p_{n_{\gamma(u)}}$ . For  $1 \leq u \leq v$ , denote  $\mathcal{P}_u$  the assertion: ‘‘there exists diagonal matrices  $\mathbf{D}_{\ell_{\gamma(1)}}, \dots, \mathbf{D}_{\ell_{\gamma(u)}} \in \mathcal{D}_r$  with nonzero diagonal entries such that, for each  $1 \leq \eta \leq u$ , we have  $\mathbf{H}'^{[q_{\gamma(\eta)};\ell_{\gamma(\eta)+1}]} = \mathbf{D}_{q_{\gamma(\eta)}}^{-1} \mathbf{H}^{[q_{\gamma(\eta)};\ell_{\gamma(\eta)+1}]} \mathbf{D}_{\ell_{\gamma(\eta)}}$  and  $(\mathbf{H}'^{[\ell_{\gamma(\eta)};p_{\gamma(\eta)}]})^\top = \mathbf{D}_{\ell_{\gamma(\eta)}}^{-1} (\mathbf{H}^{[\ell_{\gamma(\eta)};p_{\gamma(\eta)}]})^\top \mathbf{D}_{p_{\gamma(\eta)}}$ ’’, with the convention  $\mathbf{D}_L = \mathbf{I}_{2^L}$ . Then, by applying Lemma 3.12, we show inductively that  $\mathcal{P}_u$  is true for all  $1 \leq u \leq v$ . In particular,  $\mathcal{P}_v$  yields the claim of Theorem 3.8.  $\blacksquare$

*Remark 3.13.* In the case where the observed matrix  $\mathbf{Z}$  is noisy, one can adapt Line 8 of Algorithm 3.1 and set  $\mathbf{C}^i$  as the best rank-one approximation of  $(\mathbf{Z}_{k,l})_{(k,l) \in \mathcal{S}^i}$ , via singular value decomposition for instance.

**4. Conclusion.** This paper proposed an illustration of the application of the general framework introduced in the companion paper [14] to show some identifiability results on some well-known matrices, like the DFT, DST and DCT matrices. We also proposed an extension of our analysis with two factors to the multilayer case via a hierarchical method. In complement to the discussions at the end of the companion paper [14], we present some other perspectives for future work.

*Hierarchical factorization method.* An extension of this work is to relax the sparsity constraints in the multilayer case to a *family* of sparsity patterns, instead of fixed-support sparsity constraints.

*Algorithm.* From the algorithmic point of view, one approach to solve sparse matrix factorization is to minimize the approximation error between the observed matrix and the product of the sparse factors under sparsity constraints [11, 5]. An alternative approach suggested by our analysis is to rely on low-rank matrix completion operations with sparsity constraints, in the spirit of Algorithm 2.1. This kind of approach based on the lifting procedure has been

considered in algorithms for blind deconvolution [1] or signal recovery from magnitude measurements [2]. However, it still remains to explore such an approach when considering sparsity constraints.

*Stability.* One can study the stability of Algorithm 3.1 under the noisy measurement case. Also, the application of Algorithm 2.1 under the noisy case might suffer from instability issues, as the propagation scheme to complete missing entries with the rank-one constraint is unstable [4], so it is necessary to adapt Algorithm 2.1 in order to allow for stable recovery of exact matrix decomposition with fixed-support.

## REFERENCES

- [1] A. AHMED, B. RECHT, AND J. ROMBERG, *Blind deconvolution using convex programming*, IEEE Transactions on Information Theory, 60 (2013), pp. 1711–1732, <https://doi.org/10.1109/TIT.2013.2294644>.
- [2] E. J. CANDÈS, T. STROHMER, AND V. VORONINSKI, *Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming*, Communications on Pure and Applied Mathematics, 66 (2013), pp. 1241–1274, <https://doi.org/10.1002/cpa.21432>.
- [3] S. CHOUDHARY AND U. MITRA, *Identifiability scaling laws in bilinear inverse problems*, arXiv preprint arXiv:1402.2637, (2014), <https://arxiv.org/abs/1402.2637>.
- [4] A. COSSE AND L. DEMANET, *Stable rank-one matrix completion is solved by the level 2 Lasserre relaxation*, Foundations of Computational Mathematics, (2020), pp. 1–50, <https://doi.org/10.1007/s10208-020-09471-y>.
- [5] T. DAO, A. GU, M. EICHHORN, A. RUDRA, AND C. RÉ, *Learning fast algorithms for linear transforms using butterfly factorizations*, in International Conference on Machine Learning, PMLR, 2019, pp. 1517–1527, <http://proceedings.mlr.press/v97/dao19a/dao19a.pdf>.
- [6] E. GRELIER, A. NOUY, AND M. CHEVREUIL, *Learning with tree-based tensor formats*, arXiv preprint arXiv:1811.04455, (2018), <https://arxiv.org/abs/1811.04455>.
- [7] F. KIRÁLY AND R. TOMIOKA, *A combinatorial algebraic approach for the identifiability of low-rank matrix completion*, arXiv preprint arXiv:1206.6470, (2012), <https://arxiv.org/abs/1206.6470>.
- [8] Q.-T. LE, *Multilayer sparse matrix factorization*, tech. report, ENS Paris Saclay, Sept. 2020, <https://hal.inria.fr/hal-03130680>.
- [9] Q.-T. LE AND R. GRIBONVAL, *Structured support exploration for multilayer sparse matrix factorization*, in ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2021, pp. 3245–3249, <https://doi.org/10.1109/ICASSP39728.2021.9414238>.
- [10] L. LE MAGOAROU, *Matrices efficaces pour le traitement du signal et l'apprentissage automatique*, PhD thesis, INSA de Rennes, 2016, <https://hal.inria.fr/tel-01412558>. Written in French.
- [11] L. LE MAGOAROU AND R. GRIBONVAL, *Flexible multilayer sparse approximations of matrices and applications*, IEEE Journal of Selected Topics in Signal Processing, 10 (2016), pp. 688–700, <https://doi.org/10.1109/JSTSP.2016.2543461>.
- [12] Y. LI, H. YANG, E. R. MARTIN, K. L. HO, AND L. YING, *Butterfly factorization*, Multiscale Modeling & Simulation, 13 (2015), pp. 714–732, <https://doi.org/10.1137/15M1007173>.
- [13] F. MALGOUYRES AND J. LANDSBERG, *On the identifiability and stable recovery of deep/multi-layer structured matrix factorization*, in 2016 IEEE Information Theory Workshop (ITW), IEEE, 2016, pp. 315–319, <https://doi.org/10.1109/ITW.2016.7606847>.
- [14] L. ZHENG, R. GRIBONVAL, AND E. RICCIETTI, *Identifiability in exact two-layer sparse matrix factorization*, tech. report, École Normale Supérieure de Lyon, 2021. Working draft.

## Appendix A. Proof of Theorem 2.15.

*Proof.* Suppose that  $\mathcal{S}$  is closable. Let  $\mathcal{C} \in \Gamma_{\mathcal{S}}^{(=)}$ , and let us show that  $\mathcal{C} \in \mathcal{U}(\Gamma_{\mathcal{S}})$ . Let  $\mathcal{C}' \in \Gamma_{\mathcal{S}}$  such that  $\mathcal{A}(\mathcal{C}) = \mathcal{A}(\mathcal{C}')$ . The goal of the proof is to show that  $\mathcal{C} \sim \mathcal{C}'$ . Denote  $\mathcal{Z} := \mathcal{A}(\mathcal{C})$ ,  $\mathcal{G} := \mathcal{G}(\mathcal{S})$  the observable bipartite graphs associated to  $\mathcal{S}$ , and  $N := N(\mathcal{G})$ . As



$\mathcal{S}$  is closable, all the rank-one supports  $\mathcal{S}^i$  ( $1 \leq i \leq r$ ) are different, so  $\mathcal{G}^i$  is not the empty graph for any  $i \in \llbracket r \rrbracket$ . The complete graph  $\mathcal{K}_{\mathcal{G}^i}$  is denoted  $\mathcal{K}_i$  for simplicity ( $1 \leq i \leq r$ ). For any integer  $n$ , denote  $\mathcal{G}_n := (b \circ a)^{(n)}(\mathcal{G})$ , where  $a$  and  $b$  are the completion operations defined in [Definition 2.9](#), and  $\mathcal{H}_n = a(\mathcal{G}_n)$ .

For  $n \in \llbracket N \rrbracket$ , denote  $P_n$  the assertion: “for all  $i \in \llbracket r \rrbracket$ , for all  $(k, l) \in E(\mathcal{G}_n^i)$ ,  $(\mathcal{C}^i)_{k,l} = (\mathcal{C}'^i)_{k,l}$ ”. Let us prove by induction that  $P_n$  is true for all  $n \in \llbracket N \rrbracket$ . Since  $N = N(\mathcal{G})$ ,  $\mathcal{G}_N = \text{cl}(\mathcal{G})$ , and by assumption,  $\text{cl}(\mathcal{G}) = (\mathcal{K}_i)_{i=1}^r$ . Therefore, proving  $P_N$  would yield  $\mathcal{C} = \mathcal{C}'$  and end the proof.

**Initialization.** Fix  $i \in \llbracket r \rrbracket$ , and  $(k, l) \in E(\mathcal{G}_0^i)$ . By definition,  $E(\mathcal{G}_0^i) = \text{obs}_i(\mathcal{S}) = \mathcal{S}^i \setminus \left( \bigcup_{j \neq i} \mathcal{S}^j \right)$ , so for  $j \neq i$ ,  $(\mathcal{C}^j)_{k,l} = 0 = (\mathcal{C}'^j)_{k,l}$ . Therefore,  $(\mathcal{C}^i)_{k,l} = \mathbf{Z}_{k,l} = (\mathcal{C}'^i)_{k,l}$ , which shows  $P_0$ .

**Inductive step.** Let  $n \in \llbracket N-1 \rrbracket$ , and suppose that  $P_n$  is verified. Denote  $\mathcal{H} := a(\mathcal{G})$ . Recall that  $\mathcal{G}_n \preceq \mathcal{H}_n \preceq \mathcal{G}_{n+1}$ .

- For any  $i$ , for any  $(k, l) \in E(\mathcal{G}_n^i)$ ,  $(\mathcal{C}^i)_{k,l} = (\mathcal{C}'^i)_{k,l}$  by assumption  $H_n$ .
- Let  $i \in \llbracket r \rrbracket$  and  $(k, l) \in E(\mathcal{H}_n^i) \setminus E(\mathcal{G}_n^i)$ . Then,  $(k, l)$  corresponds to a missing edge in a subgraph of  $\mathcal{G}_n^i$  isomorph to  $\mathcal{K}_{2,2}^-$ . For any  $(k', l')$  of this subgraph,  $(k', l') \in E(\mathcal{G}_n^i)$  by definition of the completion operation  $a$ , so  $(\mathcal{C}^i)_{k',l'} = (\mathcal{C}'^i)_{k',l'}$ . Then, by [Proposition 2.8](#), we obtain  $(\mathcal{C}^i)_{k,l} = (\mathcal{C}'^i)_{k,l}$ . This is true for any  $i \in \llbracket r \rrbracket$  and  $(k, l) \in E(\mathcal{H}_n^i)$ .
- Finally, let  $(k, l) \in E(\mathcal{H}_{n+1}^i) \setminus E(\mathcal{H}_n^i)$ . Then, by definition of the completion operation  $b$ ,  $(k, l) \in E(\mathcal{H}_n^j)$  for all  $j \neq i$  such that  $(k, l) \in \mathcal{K}_j$ . This means that  $(\mathcal{C}^j)_{k,l} = (\mathcal{C}'^j)_{k,l}$  for such  $j$ . For  $j \neq i$  such that  $(k, l) \notin \mathcal{K}_j$ , we simply have  $(\mathcal{C}^j)_{k,l} = 0 = (\mathcal{C}'^j)_{k,l}$ . Therefore, we obtain  $(\mathcal{C}^i)_{k,l} = \mathbf{Z}_{k,l} - \sum_{j \neq i} (\mathcal{C}^j)_{k,l} = \mathbf{Z}_{k,l} - \sum_{j \neq i} (\mathcal{C}'^j)_{k,l} = (\mathcal{C}'^i)_{k,l}$ , which shows  $P_{n+1}$  and ends the inductive step.  $\blacksquare$

## Appendix B. Proof of [Proposition 2.23](#).

Let us reuse some lemmas from [\[10, Chapter 7\]](#).

**Lemma B.1** ([\[10, Lemma 4, Chapter 7\]](#)). *For any  $r$ -tuple of rank-one contributions  $\mathcal{C}$ , if  $\sum_{i=1}^r \|\mathcal{C}^i\|_0 = \|\mathcal{A}(\mathcal{C})\|_0$ , then the supports  $\{\text{supp}(\mathcal{C}^i)\}_{i=1}^r$  are pairwise disjoint.*

**Corollary B.2** ([\[10, Corollary 2, Chapter 7\]](#)). *Suppose there exists an integer  $M$  such that for all  $\mathcal{S} \in \varphi(\Omega)$ ,  $\|\mathcal{S}^i\|_0 \leq M$  for each  $i \in \llbracket r \rrbracket$ . Let  $\mathcal{C} \in \Gamma_\Omega$  such that the cardinal of  $\text{supp}(\mathcal{A}(\mathcal{C}))$  is  $rM$ . Then, the supports  $\{\text{supp}(\mathcal{C}^i)\}_{i=1}^r$  are pairwise disjoint.*

*Proof.* We have  $rM = \text{card}(\text{supp}(\mathcal{A}(\mathcal{C}))) \leq \sum_{i=1}^r \|\mathcal{C}^i\|_0 \leq \sum_{i=1}^r \|\mathcal{S}^i\|_0 \leq rM$ , so all the terms of the sum are equal and we conclude by [Lemma B.1](#).  $\blacksquare$

**Proof of [Proposition 2.23](#).** By construction,  $\|\mathcal{S}^i\|_0 \leq \frac{N}{2} \times 2 = N$  for all  $i \in \llbracket r \rrbracket$ ,  $\mathcal{S} \in \varphi(\Omega)$ , and since  $\|\mathbf{Z}\| = N^2$ , the first condition of [Proposition 2.6](#) is verified, by direct application of [Corollary B.2](#). We now focus on the second condition. Let  $\mathcal{S} \in \varphi^{\text{cp}}(\Omega)$  such that  $\{\mathcal{S}^i\}_{i=1}^r$  is a partition of  $\text{supp}(\mathbf{Z})$ , and that the rank of  $(\mathbf{Z}_{k,l})_{(k,l) \in \mathcal{S}^i}$  ( $1 \leq i \leq r$ ) is at most one. Then, necessarily, the cardinal of each support  $\mathcal{S}^i$  ( $1 \leq i \leq r$ ) is at least  $N$ , so  $\mathcal{S}^i$  has exactly  $\frac{N}{2}$  nonzero rows and 2 nonzero columns. For each  $i$ , we write  $\mathcal{S}^i := \{k_1^{(i)}, \dots, k_{\frac{N}{2}}^{(i)}\} \times \{l_1^{(i)}, l_2^{(i)}\}$

the indices of the rank-one support  $\mathcal{S}^i$ . The objective of the proof is to show that all the row indices  $k_\alpha^{(i)}$  ( $\alpha \in \llbracket \frac{N}{2} \rrbracket$ ) have the same parity, and that  $l_1^{(i)} + l_2^{(i)} - 1 = N$ .

**B.1. Case of the DCT matrix.** In this paragraph, we denote  $\mathbf{Z} = \mathbf{DCT}_N$  defined by (2.18). For each  $i$ , the submatrix  $(\mathbf{Z}_{k,l})_{(k,l) \in \mathcal{S}^i}$  is of rank one, so there exists a scalar  $\lambda^{(i)}$  such that:

$$(B.1) \quad \cos \left[ \frac{\pi}{N} \left( l_1^{(i)} - \frac{1}{2} \right) (k_\alpha^{(i)} - 1) \right] = \lambda^{(i)} \cos \left[ \frac{\pi}{N} \left( l_2^{(i)} - \frac{1}{2} \right) (k_\alpha^{(i)} - 1) \right], \quad \forall \alpha \in \llbracket \frac{N}{2} \rrbracket.$$

Fix  $i \in \llbracket N \rrbracket$ . For simplicity, we omit the subscript  $(i)$  for the rest of the proof.

**B.1.1. Case 1: odd row indices.** Suppose that there exists  $\alpha$  such that  $k_\alpha = 1$ . Let us show that  $k_\alpha$  is odd for all  $\alpha \in \llbracket \frac{N}{2} \rrbracket$ , and  $l_1 + l_2 - 1 = N$ . Indeed, evaluating (B.1) at  $k_\alpha = 1$  yields  $\lambda = 1$ . Thus, for each  $\alpha \in \llbracket \frac{N}{2} \rrbracket$ :

$$\begin{aligned} \frac{\pi}{N} \left( l_1 - \frac{1}{2} \right) (k_\alpha - 1) &\equiv \frac{\pi}{N} \left( l_2 - \frac{1}{2} \right) (k_\alpha - 1) \pmod{2\pi} \\ \text{or } \frac{\pi}{N} \left( l_1 - \frac{1}{2} \right) (k_\alpha - 1) &\equiv -\frac{\pi}{N} \left( l_2 - \frac{1}{2} \right) (k_\alpha - 1) \pmod{2\pi}, \end{aligned}$$

which is equivalent to:

$$(B.2) \quad (l_1 - l_2) (k_\alpha - 1) \equiv 0 \pmod{2N} \quad \text{or} \quad (l_1 + l_2 - 1) (k_\alpha - 1) \equiv 0 \pmod{2N}.$$

Suppose by contradiction that there exists  $\alpha$  such that  $k_\alpha$  is even. Then,  $(k_\alpha - 1)$  is odd, so by (B.2) and Euclid's lemma,  $2N$  divides either  $(l_1 - l_2)$  or  $(l_1 + l_2 - 1)$ , because  $2N$  is even. But both cases are not possible, because  $(l_1 - l_2) \in \{-N + 1, \dots, -1\} \cup \{1, \dots, N - 1\}$ , and  $(l_1 + l_2 - 1) \in \{1, \dots, 2N - 1\}$ . Therefore, all the indices  $k_\alpha$  ( $\alpha \in \llbracket \frac{N}{2} \rrbracket$ ) are odd. Suppose now by contradiction that  $l_1 + l_2 - 1 \neq N$ . Then, evaluating (B.2) for  $k_\alpha = 3$ ,  $N$  divides either  $(l_1 - l_2)$  or  $(l_1 + l_2 - 1)$ . Again, this is not possible, which shows that necessarily,  $l_1 + l_2 - 1 = N$ .

**B.1.2. Case 2: even row indices.** Suppose now there does not exist  $\alpha$  such that  $k_\alpha = 1$ . Then, by the previous case, all the column indices  $k_\alpha$  ( $\alpha \in \llbracket \frac{N}{2} \rrbracket$ ) are odd, because the supports  $\{\mathcal{S}^{i'}\}_{i'=1}^r$  are pairwise disjoint. Let us show that  $l_1 + l_2 - 1 = N$ . Evaluating (B.1) at  $k_\alpha = 2$  yields:

$$\lambda = \frac{\cos \left[ \frac{\pi}{N} \left( l_1 - \frac{1}{2} \right) \right]}{\cos \left[ \frac{\pi}{N} \left( l_2 - \frac{1}{2} \right) \right]}.$$

Then, evaluating (B.1) at  $k_\alpha = N$ :

$$\cos(aN - a) \cos(b) = \cos(a) \cos(bN - b)$$

where we denoted  $a = \frac{\pi}{N} \left( l_1 - \frac{1}{2} \right)$  and  $b = \frac{\pi}{N} \left( l_2 - \frac{1}{2} \right)$ . This is equivalent to:

$$(B.3) \quad \cos(aN - a - b) + \cos(aN - a + b) = \cos(a - bN + b) + \cos(a + bN - b).$$

We compute:

$$(B.4) \quad \begin{aligned} \cos(aN - a - b) &= \begin{cases} \sin \left[ \frac{\pi}{N}(l_1 + l_2 - 1) \right] & \text{if } l_1 \text{ is odd} \\ -\sin \left[ \frac{\pi}{N}(l_1 + l_2 - 1) \right] & \text{otherwise} \end{cases}, \\ \cos(aN - a + b) &= \begin{cases} \sin \left[ \frac{\pi}{N}(l_2 - l_1) \right] & \text{if } l_1 \text{ is odd} \\ -\sin \left[ \frac{\pi}{N}(l_2 - l_1) \right] & \text{otherwise} \end{cases}, \\ \cos(a - bN + b) &= \begin{cases} \sin \left[ \frac{\pi}{N}(l_1 + l_2 - 1) \right] & \text{if } l_2 \text{ is odd} \\ -\sin \left[ \frac{\pi}{N}(l_1 + l_2 - 1) \right] & \text{otherwise} \end{cases}, \\ \cos(a + bN - b) &= \begin{cases} \sin \left[ \frac{\pi}{N}(l_1 - l_2) \right] & \text{if } l_2 \text{ is odd} \\ -\sin \left[ \frac{\pi}{N}(l_1 - l_2) \right] & \text{otherwise} \end{cases}. \end{aligned}$$

Suppose by contradiction that  $l_1$  and  $l_2$  have the same parity. Then (B.3) combined with (B.4) yields  $\sin \left[ \frac{\pi}{N}(l_2 - l_1) \right] = 0$ , which means that  $N$  divides  $l_1 - l_2$ . This is impossible as  $(l_1 - l_2) \in \{-N + 1, \dots, -1\} \cup \{1, \dots, N - 1\}$ . Therefore,  $l_1$  and  $l_2$  do not have the same parity, and (B.3) combined with (B.4) yields  $\sin \left[ \frac{\pi}{N}(l_1 + l_2 - 1) \right] = 0$ . Therefore,  $N$  divides  $(l_1 + l_2 - 1)$ , and since  $(l_1 + l_2 - 1) \in \{1, \dots, 2N - 1\}$ , we obtain  $l_1 + l_2 - 1 = N$ .

**B.2. Case of the DST matrix.** In this paragraph, we denote  $\mathbf{Z} = \mathbf{DST}_N$  defined by (2.19). For each  $i$ , the submatrix  $(\mathbf{Z}_{k,l})_{(k,l) \in \mathcal{S}^i}$  is of rank one, so there exists a scalar  $\lambda^{(i)}$  such that:

$$(B.5) \quad \sin \left[ \frac{\pi}{N} \left( l_1^{(i)} - \frac{1}{2} \right) k_\alpha^{(i)} \right] = \lambda^{(i)} \sin \left[ \frac{\pi}{N} \left( l_2^{(i)} - \frac{1}{2} \right) k_\alpha^{(i)} \right], \quad \forall \alpha \in \left\lfloor \frac{N}{2} \right\rfloor.$$

Fix  $i \in \llbracket N \rrbracket$ . For simplicity, we omit the subscript  $(i)$  for the rest of the proof.

**B.2.1. Case 1: even row indices.** Suppose that there exists  $\alpha$  such that  $k_\alpha = N$ . Let us show that  $k_\alpha$  is even for all  $\alpha \in \left\lfloor \frac{N}{2} \right\rfloor$ , and  $l_1 + l_2 - 1 = N$ . Indeed, evaluating (B.5) at  $k_\alpha = N$  yields:

$$\cos(\pi l_1) = \lambda \cos(\pi l_2).$$

*When  $l_1$  and  $l_2$  have the same parity.* In this case,  $\lambda = 1$ , and (B.5) yields for all  $\alpha \in \left\lfloor \frac{N}{2} \right\rfloor$ :

$$\begin{aligned} \frac{\pi}{N} \left( l_1 - \frac{1}{2} \right) k_\alpha &\equiv \frac{\pi}{N} \left( l_2 - \frac{1}{2} \right) k_\alpha \pmod{2\pi}, \\ \text{or } \frac{\pi}{N} \left( l_1 - \frac{1}{2} \right) k_\alpha &\equiv \pi - \frac{\pi}{N} \left( l_2 - \frac{1}{2} \right) k_\alpha \pmod{2\pi}, \end{aligned}$$

which simplifies into:

$$(B.6) \quad (l_1 - l_2)k_\alpha \equiv 0 \pmod{2N} \quad \text{or} \quad (l_1 + l_2 - 1)k_\alpha \equiv N \pmod{2N}, \quad \forall \alpha \in \left\lfloor \frac{N}{2} \right\rfloor.$$

Suppose by contradiction that there exists  $\alpha$  such that  $k_\alpha$  is odd. Since  $(l_1 - l_2) \in \{-N + 1, \dots, -1\} \cup \{1, \dots, N - 1\}$ ,  $2N$  does not divide  $(l_1 - l_2)$ . Consequently,  $2N$  cannot divide

$(l_1 - l_2)k_\alpha$ , because  $2N$  and  $k_\alpha$  do not have the same parity. Therefore, by (B.6),  $2N$  divides necessarily  $(l_1 + l_2 - 1)k_\alpha - N$ . But  $l_1$  and  $l_2$  have the same parity, so  $(l_1 + l_2 - 1)$  is odd. As  $k_\alpha$  is odd and  $N$  is even, the number  $(l_1 + l_2 - 1)k_\alpha - N$  is odd, which is not possible. Therefore, we conclude that  $k_\alpha$  is even for all  $\alpha \in \llbracket \frac{N}{2} \rrbracket$ . Suppose now that by contradiction  $l_1 + l_2 - 1 \neq N$ . By evaluating (B.6) at  $k_\alpha = 2$ , we obtain  $(l_1 - l_2) \equiv 0 \pmod{N}$  or  $(l_1 + l_2 - 1) \equiv \frac{N}{2} \pmod{N}$ . The former is not possible, because  $(l_1 - l_2) \in \{-N + 1, \dots, -1\} \cup \{1, \dots, N - 1\}$ . The latter is also not possible, because  $(l_1 + l_2 - 1)$  is odd,  $\frac{N}{2}$  is even, and  $N$  is even. In conclusion,  $l_1 + l_2 - 1 = N$ .

*When  $l_1$  and  $l_2$  do not have the same parity.* In this case,  $\lambda = -1$ , and (B.5) yields for all  $\alpha \in \llbracket \frac{N}{2} \rrbracket$ :

$$\begin{aligned} \frac{\pi}{N} \left( l_1 - \frac{1}{2} \right) k_\alpha &\equiv -\frac{\pi}{N} \left( l_2 - \frac{1}{2} \right) k_\alpha \pmod{2\pi}, \\ \text{or } \frac{\pi}{N} \left( l_1 - \frac{1}{2} \right) k_\alpha &\equiv \pi + \frac{\pi}{N} \left( l_2 - \frac{1}{2} \right) k_\alpha \pmod{2\pi}, \end{aligned}$$

which simplifies into:

$$(B.7) \quad (l_1 + l_2 - 1)k_\alpha \equiv 0 \pmod{2N} \quad \text{or} \quad (l_1 - l_2)k_\alpha \equiv N \pmod{2N}, \quad \forall \alpha \in \llbracket \frac{N}{2} \rrbracket.$$

Suppose by contradiction that there exists  $\alpha$  such that  $k_\alpha$  is odd. Since  $(l_1 + l_2 - 1) \in \{1, \dots, 2N - 1\}$ ,  $2N$  does not divide  $(l_1 + l_2 - 1)$ . Hence,  $2N$  cannot divide  $(l_1 - l_2)k_\alpha$ , because  $2N$  and  $k_\alpha$  do not have the same parity. Therefore, by (B.6),  $2N$  divides necessarily  $(l_1 - l_2)k_\alpha - N$ . But  $l_1$  and  $l_2$  do not have the same parity, so  $(l_1 - l_2)$  is odd. As  $k_\alpha$  is odd and  $N$  is even, the number  $(l_1 - l_2)k_\alpha - N$  is odd, which is not possible. Therefore, we conclude that  $k_\alpha$  are even for all  $\alpha \in \llbracket \frac{N}{2} \rrbracket$ . Suppose now that by contradiction  $l_1 + l_2 - 1 \neq N$ . By evaluating (B.6) at  $k_\alpha = 2$ , we obtain  $(l_1 + l_2 - 1) \equiv 0 \pmod{N}$  or  $(l_1 - l_2) \equiv \frac{N}{2} \pmod{N}$ . The former is not possible, because  $(l_1 + l_2 - 1) \in \{1, \dots, N - 1\} \cup \{N + 1, \dots, 2N - 1\}$ . The latter is also not possible, because  $(l_1 - l_2)$  is odd,  $\frac{N}{2}$  is even, and  $N$  is even. In conclusion,  $l_1 + l_2 - 1 = N$ .

**B.2.2. Case 2: even row indices.** Suppose now that there does not exist  $\alpha$  such that  $k_\alpha = N$ . Then, by the previous case, all the column indices  $k_\alpha$  ( $\alpha \in \llbracket \frac{N}{2} \rrbracket$ ) are even, because the supports  $\{\mathcal{S}^{i'}\}_{i'=1}^r$  are pairwise disjoint. Let us show that  $l_1 + l_2 - 1 = N$ . Evaluating (B.5) at  $k_\alpha = 0$  yields:

$$\lambda = \frac{\sin \left[ \frac{\pi}{N} \left( l_1 - \frac{1}{2} \right) \right]}{\sin \left[ \frac{\pi}{N} \left( l_2 - \frac{1}{2} \right) \right]}.$$

Then, evaluating (B.5) at  $k_\alpha = N - 2$ :

$$\sin(aN - a) \sin(b) = \sin(a) \sin(bN - b)$$

where we denoted  $a = \frac{\pi}{N} \left( l_1 - \frac{1}{2} \right)$  and  $b = \frac{\pi}{N} \left( l_2 - \frac{1}{2} \right)$ . This is equivalent to:

$$(B.8) \quad \cos(aN - a - b) - \cos(aN - a + b) = \cos(a - bN + b) - \cos(a + bN - b). \quad \blacksquare$$

Suppose by contradiction that  $l_1$  and  $l_2$  have the same parity. Then (B.8) combined with (B.4) yields  $\sin \left[ \frac{\pi}{N} (l_2 - l_1) \right] = 0$ , which means that  $N$  divides  $l_1 - l_2$ . This is impossible as  $(l_1 - l_2) \in \{-N + 1, \dots, -1\} \cup \{1, \dots, N - 1\}$ . Therefore,  $l_1$  and  $l_2$  do not have the same parity, and (B.8) combined with (B.4) yields  $\sin \left[ \frac{\pi}{N} (l_1 + l_2 - 1) \right] = 0$ . Therefore,  $N$  divides  $(l_1 + l_2 - 1)$ , and since  $(l_1 + l_2 - 1) \in \{1, \dots, 2N - 1\}$ , we obtain  $l_1 + l_2 - 1 = N$ .