



**HAL**  
open science

# Stable recovery of the factors from a deep matrix product and application to convolutional network

François Malgouyres, Joseph Landsberg

► **To cite this version:**

François Malgouyres, Joseph Landsberg. Stable recovery of the factors from a deep matrix product and application to convolutional network. 2017. hal-01494267v1

**HAL Id: hal-01494267**

**<https://hal.science/hal-01494267v1>**

Preprint submitted on 23 Mar 2017 (v1), last revised 1 Feb 2023 (v4)

**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.

# Stable recovery of the factors from a deep matrix product and application to convolutional network

François Malgouyres<sup>1</sup> and Joseph Landsberg<sup>2</sup>

<sup>1</sup>Institut de Mathématiques de Toulouse ; UMR5219, Université de Toulouse ; CNRS,  
UPS IMT, F-31062 Toulouse Cedex 9, France

<sup>2</sup>Department of Mathematics , Mailstop 3368, Texas A& M University, College  
Station, TX 77843-3368

March 23, 2017

## Abstract

We study a deep matrix factorization problem. It takes as input a matrix  $X$  obtained by multiplying  $K$  matrices (called factors). Each factor is obtained by applying a fixed linear operator to a short vector of parameters satisfying a model (for instance sparsity, grouped sparsity, non-negativity, constraints defining a convolution network...). We call the problem deep or multi-layer because the number of factors is not limited. In the practical situations we have in mind, we can typically have  $K = 10$  or  $100$ . This work aims at identifying conditions on the structure of the model that guarantees the stable recovery of the factors from the knowledge of  $X$  and the model for the factors.

We provide necessary and sufficient conditions for the identifiability of the factors (up to a scale rearrangement). We also provide a necessary and sufficient condition called deep-Null Space Property (because of the analogy with the usual Null Space Property in the compressed sensing framework) which guarantees that even an inaccurate optimization algorithm for the factorization stably recovers the factors.

We illustrate the theory with a practical example where the deep factorization is a convolutional network.

## 1 Introduction

### 1.1 Content of the paper

We consider the following matrix factorization problem: let  $K \in \mathbb{N}$ ,  $m_1 \dots m_{K+1} \in \mathbb{N}$ , write  $m_1 = m$ ,  $m_{K+1} = n$ . We are given a matrix  $X \in \mathbb{R}^{m \times n}$  which is (approximately) the product of factors  $X_k \in \mathbb{R}^{m_k \times m_{k+1}}$ :

$$X = X_1 \cdots X_K.$$

This paper investigates models/constraints imposed on the factors  $X_k$  for which we can (up to obvious scale rearrangement) identify or stably recover the factors  $X_k$  from  $X$ .

This question is of paramount importance in many fields including statistics and machine learning, vision, image processing, information theory and numerical linear algebra. In practical applications  $X$  contains some data or represents a linear operator. It is often only specified indirectly and/or approximately. Notice that  $X$  might be a simple vector.

We now describe the structures imposed on the factors that we investigate in this paper. The factors are required to be structured matrices defined by a number  $S \in \mathbb{N}$  of unknown parameters. More precisely, for  $k = 1 \dots K$ , let

$$\begin{aligned} M_k : \mathbb{R}^S &\longrightarrow \mathbb{R}^{m_k \times m_{k+1}}, \\ h &\longmapsto M_k(h) \end{aligned} \quad (1)$$

be a linear map. This linear map might for instance simply map the values in  $h$  to prescribed locations in the matrix  $M_k(h)$ . In that case, when  $S$  is small, all the factors are  $S$ -sparse. Another insightful example is when  $M_k$  uses  $h$  to construct a Toeplitz or Circulant matrix. More complex examples include those where the matrix  $M_k(h)$  is obtained by combining several smaller Toeplitz or circulant matrices. In the latter case, the product  $M_1(h_1) \dots M_K(h_K)$  can be the matrix corresponding to a convolutional network. This example is presented in Section 7. Another interesting example is obtained when  $M_K$  (or  $M_1$ ) is defined by the product  $M'_K D$  (or  $D M'_1$ ), where the columns of  $D$  contain learning samples and  $M'_K$  (or  $M'_1$ ) has the form (1). In this case,  $X$  contains an analysis of the samples.

In addition to the structure induced by the operators  $M_k$ , we also consider structure imposed on the vectors  $h$ . We assume that we know a collection of models  $\mathcal{M} = (\mathcal{M}^L)_{L \in \mathbb{N}}$  with the property that for every  $L$ ,  $\mathcal{M}^L \subset \mathbb{R}^{S \times K}$  is a given subset. We will assume that the parameters  $\mathbf{h} \in \mathbb{R}^{S \times K}$  defining the factors are such that there exists  $L \in \mathbb{N}$  such that  $\mathbf{h} \in \mathcal{M}^L$ . The typical examples we have in mind include models for which  $\mathcal{M}^L$  is the level set of a function. The function defines a prior on the parameters  $\mathbf{h}$ . For instance, when  $K = 2$ ,  $\mathcal{M}^L$  might contain sparse vectors, impose grouped sparsity or co-sparsity. Still when  $K = 2$ , other examples impose non-negativity constraints, orthogonality (for low rank approximation), equality (in phase retrieval), an upper-bound on the norm of the columns of  $M_1(h_1)$  (in dictionary learning). The examples are numerous.

We now precisely state the problem considered in this paper. We assume a collection of models  $\mathcal{M}$  is known and that there exists a model  $\mathcal{M}^{\bar{L}}$  defined by an unknown  $\bar{L}$  and parameters  $(\bar{h}_k)_{k=1..K} \in \mathcal{M}^{\bar{L}}$ , with  $\bar{h}_k \in \mathbb{R}^S$  for all  $k = 1..K$ , and we only are given the product of matrices

$$X = M_1(\bar{h}_1) \dots M_K(\bar{h}_K) + e,$$

for an unknown error term  $e$ . Our goal is to establish sharp and, as far as possible, simple conditions guaranteeing that we can recover the parameters  $(\bar{h}_k)_{k=1..K}$  with an accuracy comparable to  $\|e\|$ . The typical solver we have in mind minimizes

$$\operatorname{argmin}_{L \in \mathbb{N}, (h_k)_{k=1..K} \in \mathcal{M}^L} \|M_1(h_1) \dots M_K(h_K) - X\|^2. \quad (2)$$

However, following [7], we state stability results that hold for an inaccurate (also called inexact or approximate) minimizer of (2). Such an inaccurate minimizer  $(h_k)_{k=1..K}$  is such that  $(h_k)_{k=1..K} \in \mathcal{M}^L$ , for some  $L \in \mathbb{N}$ , and

$$\|M_1(h_1) \dots M_K(h_K) - X\|^2 \text{ is small.}$$

Doing so, the stability results apply to a larger class of solvers. In particular, it can apply to heuristics or (for instance) to solvers that minimize an approximate objective function or an objective function that is not based on the Euclidean distance. Our statements say that the

smaller  $\|M_1(h_1) \dots M_K(h_K) - X\|^2$  the smaller the error on the estimate of the parameters  $\bar{h}_k$ .

The stability question boils down to an identifiability issue when the error term  $e$  vanishes and the minimization is exact. We therefore begin to study necessary and sufficient conditions for the uniqueness of the factors satisfying both  $(\bar{h}_k)_{k=1..K} \in \mathcal{M}^L$  and

$$M_1(\bar{h}_1) \dots M_K(\bar{h}_K) = X. \quad (3)$$

The minimization problem (2) is non-convex because the product  $M_1(h_1) \dots M_K(h_K)$  is not (jointly) linear. The constraint  $\cup_{L \in \mathbb{N}} \mathcal{M}^L$  might also be non-convex. As a consequence, solving or even finding an efficient heuristic solving (2) might be difficult or impossible for some instances of the problem. We do not address the numerical issues related to the minimization of (2). In this regard, although the identifiability is desired when interpreting the solution, it implies that the minimizer of (2) is unique. Intuitively, this is expected to reduce the size of the convergence basin and complicate the numerical resolution of (2). In that sense, a sharp condition of identifiability separates identifiable problems and problems which better lend themselves to global optimization. Outside of this crude intuition and the fact that, we do not investigate whether (2) can actually be minimized or not (see [21] for an example of such a result).

When  $K \geq 2$ , (3) and (2) are in general highly non-linear/non-convex (even when  $\bar{L}$  is known and  $\mathcal{M}^{\bar{L}}$  is a Euclidean space), so the uniqueness and stability of the solution is not easy to characterize. This is due to the product of the factors. However, our results also apply when  $K = 1$ . When  $K = 1$ , the questions and results stated in this work are identical or close to well established existing results in compressed/compressive sensing [8, 15, 7, 13]. Although it is not our primary interest, for pedagogical reasons, we also express our definitions and statements in this setting.

The main contributions of this paper are:

- In the absence of noise (see Section 5):
  - We establish a simple geometric condition on the intersection of two sets which are necessary and sufficient to guarantee the identifiability of the parameters  $\bar{h}_k$  defining the factors (Proposition 7).
  - We provide simpler conditions which involve the rank of a certain linear operator (defined in Section 4) such that:
    - \* If this rank is larger than another quantity, (3) is almost surely identifiable (Theorem 3).
    - \* If this rank is below a certain constant, (3) is not identifiable (Theorem 4);
  - We also provide a simple algorithm to compute this rank that works in many reasonable cases (Proposition 4).
- In the presence of noise when considering an inaccurate minimizer (see Section 6):
  - We define the deep-Null Space Property (Definition 3): a generalization of the usual Null Space Property [13] that also applies to the deep matrix factorization problem.
  - We establish that when the deep-Null Space Property holds we can recover the factors with an accuracy bounded above by the sum of the noise level and a quantity reflecting the minimization inaccuracy (Theorem 5).
  - We establish the converse statement: if we are able to recover the factors with an accuracy upper bounded by the noise level then the deep-Null Space Property holds (Theorem 6).

- We specialize the above results to convolutional networks and establish a simple condition, that can be computed in many contexts, such that
  - If the condition is satisfied the convolutional networks can be stably recovered (see Theorem 7);
  - If the condition is not satisfied, the convolutional network is not identifiable (see Proposition 8).

In order to establish these results, we investigate and recall several results on tensors, tensor rank and the Segre embedding (see Section 3). In particular, we investigate the Lipschitz continuity (Theorem 2) and stable recovery (Theorem 1) for the Segre embedding.

## 1.2 Bibliographical landmarks

Matrix factorization problems are ubiquitous in statistics, information theory and data representation. The simplest version consists of a model with one layer (i.e.,  $K = 1$ ) and  $\mathcal{M} = \mathbb{R}^{S \times K}$ . This is the usual linear approximation problem. In this case,  $X$  can be vectorized to form a column vector and the operator  $M_1$  simply multiplies the column vector  $h_1$  by a fixed (rectangular) matrix. Typically, in this setting, the latter matrix has more rows than columns and, when a solution to (3) exists, its uniqueness depends on the column rank of the matrix.

The above linear approximation is often improved using a “non-linear approximation” [14]. In this framework, the fixed matrix has more columns than rows and  $h_1$  contains a sparse vector whose support is also optimized to better approximate the data. The identifiability and stable recovery for this problem has been intensively studied and gave rise to a new application named compressed/compressive sensing (see [8, 15]). Some compressed sensing statements (especially the ones guaranteeing that any minimizer of the  $\ell^0$  problem stably recovers the unknown) are special cases ( $K = 1$ ) of the statements provided in this paper. We will not perform a complete review on compressed sensing but would like to highlight the Null Space Property described in [13]. The fundamental limits of compressed sensing (for a solution of the  $\ell^0$  problem) have been analyzed in detail in [7].

The questions we are studying are mostly relevant (and new) when  $K \geq 2$ . In the case of such models, the non-linearity comes from the multiplicative nature of (3) and the identifiability and stable recovery are not easily guaranteed. Recently, sparse coding and dictionary learning has been introduced (see [40] for an overview on the subject). In that framework,  $X$  contains the data and (most of the time) people consider two layers:  $K = 2$ . The layer  $M_1(h_1)$  is an optimized dictionary of atoms and each column of  $M_2(h_2)$  contains the code (or coordinates) of the corresponding column in  $X$ . In this case, the mapping  $M_2$  maps a vector from a small vector space into a sparse matrix. The identifiability and stable recovery of the factors has been studied in many dictionary learning contexts and provides guarantees on the approximate recovery of both an incoherent dictionary and sparse coefficients when the number of samples is sufficiently large (i.e.,  $n$  is large, in our setting). In [20], the authors developed local optimality conditions in the noiseless case, as well as sample complexity bounds for local recovery when  $M_1(h_1)$  is square and  $M_2(h_2)$  are iid Bernoulli-Gaussian. This was extended to overcomplete dictionaries in [17] (see also [43] for tight frames) and to the noisy case in [25]. The authors of [45] provide exact recovery results for dictionary learning, when the coefficient matrix has Bernoulli-Gaussian entries and the dictionary matrix has full column rank. This was extended to overcomplete dictionaries in [1] and in [5] but only for approximate recovery. Finally, [19] provides such guarantees under general conditions which cover many practical settings.

Factorizations with  $K = 2$  have also been considered for the purpose of phase retrieval [9], blind deconvolution [2, 12, 30], blind deconvolution in random mask imaging [6], blind-deconvolution and blind-de-mixing [31], self-calibration [32] and Non-negative matrix factorization [29, 16, 28, 4]. Most of these papers use the same lifting property we are using. They further propose to convexify the problem and provide sufficient conditions for obtaining identifiability and stability. A more general bilinear framework is considered in [12], where the analysis shares similarities with the results presented here but is restricted to identifiability when  $K = 2$ .

When compared to these results, the scope of the present study is to consider the identifiability and the stability of the recovery for any  $K \geq 1$ , in a general context. The authors have announced preliminary versions of the results described here in [37]. They are significantly extended here. To the best of our knowledge, little is known concerning the identifiability and the stability of matrix factorization when  $K \geq 3$ . The uniqueness of the factorization corresponding to the Fast Fourier Transform was proved in [34]. Other results consider the identifiability of the factors which are sparse and random [39] and might even consider the presence of non-linearities between the layers to include the deep classification architectures [3].

The use of deep matrix factorization is classical. In particular handcrafted deep matrix factorization of a few particular matrices are used in many fields of mathematics and engineering. Most fast transforms, such as the Cooley-Tukey Fast Fourier Transform, the Discrete Cosine Transform and the Wavelet transform, are deep matrix products.

The construction of optimized deep matrix factorization only started recently (see [10, 11] and references therein). In [10, 11, 36], the authors consider compositions of sparse convolutions organized according to a convolutional tree. In the simplified case studied in [10],  $X$  is a vector, the vectors  $h_k$  are the convolution kernels and each operator  $M_k$  maps  $h_k$  to a circulant (or block-circulant) matrix. The first layer corresponds to the coordinates/code of  $X$  in the frame obtained by computing the compositions of convolutions along the unique branch of the tree. In [36], the authors consider a factorization involving several sparse layers. In that work, the authors simultaneously estimate the support and the coefficients of each sparse factor. They use this factorization to define an analogue of the Fast Fourier Transform for signals living on graphs [35] and latter reworked on this principle using old ideas [18]. In [26], the authors consider a (deep) multi-resolution matrix factorization, inspired by the wavelet decomposition, where the factors are orthogonal and sparse. In [41, 42], the authors consider factors based on householder reflectors and Givens rotation. In [33], the authors study a multi-layer Non-negative matrix factorization.

## 2 Notation and summary of the hypotheses

We continue to use the notation introduced in the introduction. For an integer  $k \in \mathbb{N}$ , set  $\mathbb{N}_k = \{1, \dots, k\}$ .

We consider  $K \geq 1$  and  $S \geq 2$  and real valued tensors of order  $K$  whose axes are of size  $S$ , denoted by  $T \in \mathbb{R}^{S \times \dots \times S}$ . The space of tensors is abbreviated  $\mathbb{R}^{S^K}$ . The entries of  $T$  are denoted by  $T_{i_1, \dots, i_K}$ , where  $(i_1, \dots, i_K) \in (\mathbb{N}_S)^K$ . The index set is simply denoted  $\mathbb{N}_S^K$ . For  $\mathbf{i} \in \mathbb{N}_S^K$ , the entries of  $\mathbf{i}$  are  $\mathbf{i} = (i_1, \dots, i_K)$  (for  $\mathbf{j} \in \mathbb{N}_S^K$  we let  $\mathbf{j} = (j_1, \dots, j_K)$  etc.). We either write  $T_{\mathbf{i}}$  or  $T_{i_1, \dots, i_K}$ .

A collection of vectors is denoted  $\mathbf{h} \in \mathbb{R}^{S \times K}$  (i.e., using bold fonts). Our collections are composed of  $K$  vectors of size  $S$  and the  $k^{\text{th}}$  vector is denoted  $\mathbf{h}_k \in \mathbb{R}^S$ . The  $i^{\text{th}}$  entry of the  $k^{\text{th}}$  vector is denoted  $\mathbf{h}_{k,i} \in \mathbb{R}$ . A vector not related to a collection of vectors is denoted by

$h \in \mathbb{R}^S$  (i.e., using a light font). Throughout the paper we assume

$$\mathcal{M} = (\mathcal{M}^L)_{L \in \mathbb{N}}, \text{ with } \mathcal{M}^L \subset \mathbb{R}^{S \times K}.$$

We also assume that, for all  $L \in \mathbb{N}$ ,  $\mathcal{M}^L \neq \emptyset$ . They can however be equal or constant after a given  $L'$ .

All the vector spaces  $\mathbb{R}^{S^k}$ ,  $\mathbb{R}^{S \times K}$ ,  $\mathbb{R}^S$  etc. are equipped with the usual Euclidean norm. This norm is denoted  $\|\cdot\|$  and the scalar product  $\langle \cdot, \cdot \rangle$ . In the particular case of matrices,  $\|\cdot\|$  corresponds to the Frobenius norm. We also use the usual  $p$  norm, for  $p \in [1, \infty]$ , and denote it by  $\|\cdot\|_p$ .

Set

$$\mathbb{R}_*^{S \times K} = \{\mathbf{h} \in \mathbb{R}^{S \times K}, \forall k \in \mathbb{N}_K, \|\mathbf{h}_k\| \neq 0\}. \quad (4)$$

Define an equivalence relation in  $\mathbb{R}_*^{S \times K}$ : for any  $\mathbf{h}, \mathbf{g} \in \mathbb{R}_*^{S \times K}$ ,  $\mathbf{h} \sim \mathbf{g}$  if and only if there exists  $(\lambda_k)_{k \in \mathbb{N}_K} \in \mathbb{R}^K$  such that

$$\prod_{k=1}^K \lambda_k = 1 \quad \text{and} \quad \forall k \in \mathbb{N}_K, \mathbf{h}_k = \lambda_k \mathbf{g}_k.$$

Denote the equivalence class of  $\mathbf{h} \in \mathbb{R}_*^{S \times K}$  by  $[\mathbf{h}]$ .

We say that the zero tensor is of rank 0. We say that a non-zero tensor  $T \in \mathbb{R}^{S^K}$  is of rank 1 (or decomposable) if and only if there exists a collection of vectors  $\mathbf{h} \in \mathbb{R}_*^{S \times K}$  such that  $T$  is the outer product of the vectors  $\mathbf{h}_k$ , for  $k \in \mathbb{N}_K$ , that is, for any  $\mathbf{i} \in \mathbb{N}_S^K$ ,

$$T_{\mathbf{i}} = \mathbf{h}_{1,i_1} \dots \mathbf{h}_{K,i_K}.$$

Let  $\Sigma_1 \subset \mathbb{R}^{S^K}$  denote the set of tensors of rank 0 or 1.

The rank of any tensor  $T \in \mathbb{R}^{S^K}$  is defined to be

$$\text{rk}(T) = \min\{r \in \mathbb{N} \mid \text{there exists } T_1, \dots, T_r \in \Sigma_1 \text{ such that } T = T_1 + \dots + T_r\}.$$

For  $r \in \mathbb{N}$ , let

$$\Sigma_r = \{T \in \mathbb{R}^{S^K}, \text{rk}(T) \leq r\}.$$

The \* superscript refers to optimal solutions. A set with a \* subscript means that 0 is ruled out of the set. In particular,  $\Sigma_{1,*}$  denotes the non-zero tensors of rank 1. Attention should be paid to  $\mathbb{R}_*^{S \times K}$  since its definition is not straightforward (see (4)).

### 3 Facts on the Segre embedding and tensors of rank 1 and 2

Parametrize  $\Sigma_1 \subset \mathbb{R}^{S^K}$  by the map

$$\begin{aligned} P : \mathbb{R}^{S \times K} &\longrightarrow \Sigma_1 \subset \mathbb{R}^{S^K} \\ \mathbf{h} &\longmapsto (\mathbf{h}_{1,i_1} \mathbf{h}_{2,i_2} \dots \mathbf{h}_{K,i_K})_{\mathbf{i} \in \mathbb{N}_S^K} \end{aligned} \quad (5)$$

The map  $P$  is called the Segre embedding and is often denoted by  $\widehat{\text{Seg}}$  in the algebraic geometry literature.

**Standard Facts:**

1. **Identifiability of  $[\mathbf{h}]$  from  $P(\mathbf{h})$ :** For  $\mathbf{h}$  and  $\mathbf{g} \in \mathbb{R}_*^{S \times K}$ ,  $P(\mathbf{h}) = P(\mathbf{g})$  if and only if  $[\mathbf{h}] = [\mathbf{g}]$ .
2. **Geometrical description of  $\Sigma_{1,*}$ :**  $\Sigma_{1,*}$  is a smooth (i.e.,  $C^\infty$ ) manifold of dimension  $K(S-1) + 1$  (see, e.g., [27], chapter 4, pp. 103).
3. **Geometrical description of  $\Sigma_2$ :** When  $K \geq 2$ , the singular locus  $(\Sigma_2)_{sing}$  has dimension strictly less than that of  $\Sigma_2$ , and  $\Sigma_2 \setminus (\Sigma_2)_{sing}$  is a smooth manifold. This smooth manifold is of dimension  $2K(S-1) + 2$  when  $K > 2$ , and of dimension  $4(S-1)$  when  $K = 2$  (see, e.g., [27], chapter 5).

We can improve Standard Fact 1 and obtain a stability result guaranteeing, that if we know a rank 1 tensor sufficiently close to  $P(\mathbf{h})$ , we approximately know  $[\mathbf{h}]$ . In order to state this, we need to define a metric on  $\mathbb{R}_*^{S \times K} / \sim$ . This has to be considered with care since, whatever  $\mathbf{h} \in \mathbb{R}_*^{S \times K}$ , the subset  $\{h \mid h \in [\mathbf{h}]\}$  is not compact. In particular, considering

$$\mathbf{h}'_k = \begin{cases} \lambda \mathbf{h}_k & , \text{if } k = 1 \\ \lambda^{-\frac{1}{K}} \mathbf{h}_k & , \text{otherwise} \end{cases}$$

when  $\lambda$  goes to infinity, we easily construct examples that make the standard metric on equivalence classes useless<sup>1</sup>.

This leads us to consider

$$\mathbb{R}_{diag}^{S \times K} = \{\mathbf{h} \in \mathbb{R}_*^{S \times K}, \forall k \in \mathbb{N}_K, \|\mathbf{h}_k\|_\infty = \|\mathbf{h}_1\|_\infty\}.$$

The interest in this set comes from the fact that, whatever  $\mathbf{h} \in \mathbb{R}_*^{S \times K}$ , the set  $[\mathbf{h}] \cap \mathbb{R}_{diag}^{S \times K}$  is finite. Indeed, if  $\mathbf{g} \in [\mathbf{h}] \cap \mathbb{R}_{diag}^{S \times K}$  the  $(\lambda_k)_{k \in \mathbb{N}_K} \in \mathbb{R}^K$  such that, for all  $k \in \mathbb{N}_K$ ,  $\mathbf{h}_k = \lambda_k \mathbf{g}_k$  must all satisfy  $|\lambda_k| = 1$ , i.e.  $\lambda_k = \pm 1$ .

**Definition 1.** For any  $p \in [1, \infty]$ , we define the mapping  $d_p : (\mathbb{R}_*^{S \times K} / \sim \times \mathbb{R}_*^{S \times K} / \sim) \rightarrow \mathbb{R}$  by

$$d_p([\mathbf{h}], [\mathbf{g}]) = \inf_{\substack{\mathbf{h}' \in [\mathbf{h}] \cap \mathbb{R}_{diag}^{S \times K} \\ \mathbf{g}' \in [\mathbf{g}] \cap \mathbb{R}_{diag}^{S \times K}}} \|\mathbf{h}' - \mathbf{g}'\|_p, \quad \forall \mathbf{h}, \mathbf{g} \in \mathbb{R}_*^{S \times K}.$$

**Proposition 1.** For any  $p \in [1, \infty]$ ,  $d_p$  is a metric on  $\mathbb{R}_*^{S \times K} / \sim$ .

*Proof.* Notice that, the sets  $[\mathbf{h}] \cap \mathbb{R}_{diag}^{S \times K}$  and  $[\mathbf{g}] \cap \mathbb{R}_{diag}^{S \times K}$  are finite and therefore the infimum in the definition of  $d$  is reached. We also have whatever  $\mathbf{h}, \mathbf{g} \in \mathbb{R}_*^{S \times K}$

$$d_p([\mathbf{h}], [\mathbf{g}]) = \inf_{\mathbf{h}' \in [\mathbf{h}] \cap \mathbb{R}_{diag}^{S \times K}} \left( \inf_{\mathbf{g}' \in [\mathbf{g}] \cap \mathbb{R}_{diag}^{S \times K}} \|\mathbf{h}' - \mathbf{g}'\|_p \right). \quad (6)$$

---

<sup>1</sup>For instance, if  $\mathbf{h}$  and  $\mathbf{g} \in \mathbb{R}_*^{S \times K}$  are such that (for instance)  $\mathbf{h}_1 = \mathbf{g}_1$ , we have

$$\inf_{\mathbf{h}' \in [\mathbf{h}], \mathbf{g}' \in [\mathbf{g}]} \|\mathbf{h}' - \mathbf{g}'\|_p = 0$$

even though we might have  $\mathbf{h}_2 \neq \mathbf{g}_2$  (and therefore  $[\mathbf{h}] \neq [\mathbf{g}]$ ). This does not define a metric.

Also, when  $\mathbf{h}$  and  $\mathbf{g}$  are such that  $\mathbf{h}_k \neq \mathbf{g}_k$ , whatever  $k \in \mathbb{N}_K$ , we have

$$\sup_{\mathbf{h}' \in [\mathbf{h}]} \inf_{\mathbf{g}' \in [\mathbf{g}]} \|\mathbf{h}' - \mathbf{g}'\|_p = +\infty.$$

Therefore, the Hausdorff distance between  $[\mathbf{h}]$  and  $[\mathbf{g}]$  is infinite for almost every pair  $(\mathbf{h}, \mathbf{g})$ . This metric is therefore not very useful in the present context.



Moreover, whatever  $\mathbf{h} \in \mathbb{R}_*^{S \times K}$  and  $\mathbf{h}'$  and  $\mathbf{h}'' \in [\mathbf{h}] \cap \mathbb{R}_{\text{diag}}^{S \times K}$  there exist  $(s_k)_{k \in \mathbb{N}_K} \in \{-1, 1\}^K$  such that  $\prod_{k \in \mathbb{N}_K} s_k = 1$  and

$$\mathbf{h}'_k = s_k \mathbf{h}''_k, \quad \forall k \in \mathbb{N}_K.$$

Using the above two properties, we can check that

$$\inf_{\mathbf{g}' \in [\mathbf{g}] \cap \mathbb{R}_{\text{diag}}^{S \times K}} \|\mathbf{h}' - \mathbf{g}'\|_p = \inf_{\mathbf{g}' \in [\mathbf{g}] \cap \mathbb{R}_{\text{diag}}^{S \times K}} \|\mathbf{h}'' - \mathbf{g}'\|_p$$

As a consequence, the outer infimum in (6) is irrelevant and we have

$$d_p([\mathbf{h}], [\mathbf{g}]) = \inf_{\mathbf{g}' \in [\mathbf{g}] \cap \mathbb{R}_{\text{diag}}^{S \times K}} \|\mathbf{h}' - \mathbf{g}'\|_p, \quad \forall \mathbf{h}, \mathbf{g} \in \mathbb{R}_*^{S \times K} \text{ and } \mathbf{h}' \in [\mathbf{h}] \cap \mathbb{R}_{\text{diag}}^{S \times K}.$$

Using this last property, we easily check that  $d_p$  is a metric on  $\mathbb{R}_*^{S \times K} / \sim$ .  $\square$

Using this metric, we can state, not only  $[\mathbf{h}]$  is uniquely determined by  $P(\mathbf{h})$ , but this operation is stable.

**Theorem 1. Stability of  $[\mathbf{h}]$  from  $P(\mathbf{h})$**

Let  $\mathbf{h}$  and  $\mathbf{g} \in \mathbb{R}_*^{S \times K}$  be such that  $\|P(\mathbf{g}) - P(\mathbf{h})\|_\infty \leq \frac{1}{2} \max(\|P(\mathbf{h})\|_\infty, \|P(\mathbf{g})\|_\infty)$ . For all  $p, q \in [1, \infty]$ ,

$$d_p([\mathbf{h}], [\mathbf{g}]) \leq 7(KS)^{\frac{1}{p}} \min\left(\|P(\mathbf{h})\|_\infty^{\frac{1}{K}-1}, \|P(\mathbf{g})\|_\infty^{\frac{1}{K}-1}\right) \|P(\mathbf{h}) - P(\mathbf{g})\|_q. \quad (7)$$

*Proof.* Notice first that when  $K = 1$  the inequality is a straightforward consequence of the usual inequalities between  $l^p$  norms. We therefore assume from now on that  $K \geq 2$ .

All along the proof, we consider  $\mathbf{h}$  and  $\mathbf{g} \in \mathbb{R}_*^{S \times K}$  and assume that  $\|P(\mathbf{h})\|_\infty \geq \|P(\mathbf{g})\|_\infty$ . We also assume that  $\|P(\mathbf{g}) - P(\mathbf{h})\|_\infty \leq \frac{1}{2} \|P(\mathbf{h})\|_\infty$ . We first prove the inequality when  $p = q = +\infty$ .

In order to do so, we consider

$$\mathbf{i} \in \operatorname{argmax}_{\mathbf{j} \in \mathbb{N}_S^K} |P(\mathbf{h})_{\mathbf{j}}|$$

and assume, without loss of generality (otherwise, we can multiply one vector of  $\mathbf{h}$  and  $\mathbf{g}$  by  $-1$  to get this property and multiply back once the inequality have been established), that  $P(\mathbf{h})_{\mathbf{i}} \geq 0$ . We therefore have  $P(\mathbf{h})_{\mathbf{i}} = \|P(\mathbf{h})\|_\infty$ . Notice also that we have, under the above hypotheses,

$$\|P(\mathbf{g})\|_\infty \geq P(\mathbf{g})_{\mathbf{i}} \geq P(\mathbf{h})_{\mathbf{i}} - \|P(\mathbf{g}) - P(\mathbf{h})\|_\infty \geq \frac{1}{2} \|P(\mathbf{h})\|_\infty > 0. \quad (8)$$

Moreover, we consider the operator  $E_{\mathbf{i}}$  that extracts the  $K$  signals of size  $S$  that are obtained when freezing, at the index  $\mathbf{i}$  in a tensor  $T$ , all coordinates but one. Formally, we denote

$$\begin{aligned} E_{\mathbf{i}} : \mathbb{R}^{S^K} &\longrightarrow \mathbb{R}^{S \times K} \\ T &\longmapsto E_{\mathbf{i}}(T) \end{aligned}$$

where for all  $k \in \mathbb{N}_K$  and all  $j \in \mathbb{N}_S$

$$E_{\mathbf{i}}(T)_{k,j} = T_{\mathbf{i}_1, \dots, \mathbf{i}_{k-1}, j, \mathbf{i}_{k+1}, \dots, \mathbf{i}_K}.$$

We consider

$$\mathbf{h}' = (P(\mathbf{h})_{\mathbf{i}})^{-1+\frac{1}{K}} E_{\mathbf{i}}(P(\mathbf{h})) \quad \text{and} \quad \mathbf{g}' = (P(\mathbf{g})_{\mathbf{i}})^{-1+\frac{1}{K}} E_{\mathbf{i}}(P(\mathbf{g})).$$

We have for all  $\mathbf{j} \in \mathbb{N}_S^K$

$$\begin{aligned} P(\mathbf{h}')_{\mathbf{j}} &= (P(\mathbf{h})_{\mathbf{i}})^{-K+1} P(E_{\mathbf{i}}(P(\mathbf{h})))_{\mathbf{j}}, \\ &= (P(\mathbf{h})_{\mathbf{i}})^{-K+1} \prod_{k=1}^K P(\mathbf{h})_{\mathbf{i}_1, \dots, \mathbf{i}_{k-1}, \mathbf{j}_k, \mathbf{i}_{k+1}, \dots, \mathbf{i}_K} \\ &= (P(\mathbf{h})_{\mathbf{i}})^{-K+1} \prod_{k=1}^K \mathbf{h}_{1, \mathbf{i}_1} \dots \mathbf{h}_{k-1, \mathbf{i}_{k-1}} \mathbf{h}_{k, \mathbf{j}_k} \mathbf{h}_{k+1, \mathbf{i}_{k+1}} \dots \mathbf{h}_{K, \mathbf{i}_K} \\ &= \prod_{k=1}^K \mathbf{h}_{k, \mathbf{j}_k} = P(\mathbf{h})_{\mathbf{j}}. \end{aligned}$$

We therefore have  $P(\mathbf{h}') = P(\mathbf{h})$ . This can be written  $\mathbf{h}' \in [\mathbf{h}]$ . Similarly, we have  $\mathbf{g}' \in [\mathbf{g}]$ .

Also, because of the definition of  $\mathbf{i}$  and  $\mathbf{h}'$ , we are guaranteed that, whatever  $k \in \mathbb{N}_K$ ,

$$\begin{aligned} \|\mathbf{h}'_k\|_{\infty} &= (P(\mathbf{h})_{\mathbf{i}})^{-1+\frac{1}{K}} \|E_{\mathbf{i}}(P(\mathbf{h}))\|_{\infty} \\ &= \|P(\mathbf{h})\|_{\infty}^{-1+\frac{1}{K}} \|P(\mathbf{h})\|_{\infty} = \|P(\mathbf{h})\|_{\infty}^{\frac{1}{K}} \end{aligned}$$

The latter being independent of  $k$ , we have  $\mathbf{h}' \in \mathbb{R}_{\text{diag}}^{S \times K}$ . Unfortunately, unless for instance  $\mathbf{i} \in \text{argmax}_{\mathbf{j} \in \mathbb{N}_S^K} |P(\mathbf{g})_{\mathbf{j}}|$ , it might occur that  $\mathbf{g}' \notin \mathbb{R}_{\text{diag}}^{S \times K}$ . However, if we consider

$$\mathbf{g}'' \in \text{argmin}_{\mathbf{f} \in [\mathbf{g}] \cap \mathbb{R}_{\text{diag}}^{S \times K}} \|\mathbf{f} - \mathbf{g}'\|_{\infty},$$

we have since  $\mathbf{h}' \in [\mathbf{h}] \cap \mathbb{R}_{\text{diag}}^{S \times K}$  and  $\mathbf{g}'' \in [\mathbf{g}] \cap \mathbb{R}_{\text{diag}}^{S \times K}$

$$\begin{aligned} d_{\infty}([\mathbf{h}], [\mathbf{g}]) &\leq \|\mathbf{h}' - \mathbf{g}''\|_{\infty} \\ &\leq \|\mathbf{h}' - \mathbf{g}'\|_{\infty} + \|\mathbf{g}' - \mathbf{g}''\|_{\infty}. \end{aligned} \tag{9}$$

In the sequel we will successively calculate upper bounds of  $\|\mathbf{h}' - \mathbf{g}'\|_{\infty}$  and  $\|\mathbf{g}' - \mathbf{g}''\|_{\infty}$  in order to find an upper bound of  $d_{\infty}([\mathbf{h}], [\mathbf{g}])$ .

**Upper bound of  $\|\mathbf{h}' - \mathbf{g}'\|_{\infty}$ :**

We have

$$\begin{aligned} \|\mathbf{h}' - \mathbf{g}'\|_{\infty} &= \|(P(\mathbf{h})_{\mathbf{i}})^{-1+\frac{1}{K}} E_{\mathbf{i}}(P(\mathbf{h})) - (P(\mathbf{g})_{\mathbf{i}})^{-1+\frac{1}{K}} E_{\mathbf{i}}(P(\mathbf{g}))\|_{\infty} \\ &\leq \|(P(\mathbf{h})_{\mathbf{i}})^{-1+\frac{1}{K}} (E_{\mathbf{i}}(P(\mathbf{h})) - E_{\mathbf{i}}(P(\mathbf{g})))\|_{\infty} \\ &\quad + \left\| \left( (P(\mathbf{h})_{\mathbf{i}})^{-1+\frac{1}{K}} - (P(\mathbf{g})_{\mathbf{i}})^{-1+\frac{1}{K}} \right) E_{\mathbf{i}}(P(\mathbf{g})) \right\|_{\infty} \\ &\leq \|P(\mathbf{h})\|_{\infty}^{-1+\frac{1}{K}} \|E_{\mathbf{i}}(P(\mathbf{h})) - E_{\mathbf{i}}(P(\mathbf{g}))\|_{\infty} + \|P(\mathbf{g})\|_{\infty} |(P(\mathbf{h})_{\mathbf{i}})^{-1+\frac{1}{K}} - (P(\mathbf{g})_{\mathbf{i}})^{-1+\frac{1}{K}}| \\ &\leq \|P(\mathbf{h})\|_{\infty}^{-1+\frac{1}{K}} \|P(\mathbf{h}) - P(\mathbf{g})\|_{\infty} + \|P(\mathbf{h})\|_{\infty} |(P(\mathbf{h})_{\mathbf{i}})^{-1+\frac{1}{K}} - (P(\mathbf{g})_{\mathbf{i}})^{-1+\frac{1}{K}}| \end{aligned}$$

But we also have using the mean value theorem and (8)

$$\begin{aligned}
|(P(\mathbf{h})_i)^{-1+\frac{1}{K}} - (P(\mathbf{g})_i)^{-1+\frac{1}{K}}| &\leq \left(1 - \frac{1}{K}\right) P(\mathbf{g})_i^{-2+\frac{1}{K}} |P(\mathbf{h})_i - P(\mathbf{g})_i| \\
&\leq \left(1 - \frac{1}{K}\right) \left(\frac{1}{2} \|P(\mathbf{h})\|_\infty\right)^{-2+\frac{1}{K}} \|P(\mathbf{h}) - P(\mathbf{g})\|_\infty \\
&\leq 4 \|P(\mathbf{h})\|_\infty^{-2+\frac{1}{K}} \|P(\mathbf{h}) - P(\mathbf{g})\|_\infty
\end{aligned}$$

We therefore finally obtain that

$$\|\mathbf{h}' - \mathbf{g}'\|_\infty \leq 5 \|P(\mathbf{h})\|_\infty^{-1+\frac{1}{K}} \|P(\mathbf{h}) - P(\mathbf{g})\|_\infty. \quad (10)$$

**Upper bound of  $\|\mathbf{g}' - \mathbf{g}''\|_\infty$ :**

First, since  $\mathbf{g}'' \in [\mathbf{g}] = [\mathbf{g}']$ , we know that there exists  $(\lambda_k)_{k \in \mathbb{N}_K} \in \mathbb{R}^K$  such that

$$\prod_{k=1}^K \lambda_k = 1 \quad (11)$$

and

$$\mathbf{g}''_k = \lambda_k \mathbf{g}'_k, \quad \text{for all } k \in \mathbb{N}_K.$$

Furthermore, we have for all  $k \in \mathbb{N}_K$

$$\|\mathbf{g}'_k - \mathbf{g}''_k\|_\infty = |1 - \lambda_k| \|\mathbf{g}'_k\|_\infty. \quad (12)$$

Also, if there is  $k'$  such that  $\lambda_{k'} < 0$ , since (11) holds, there necessarily exist another  $k''$  such that  $\lambda_{k''} < 0$ . If we replace  $\mathbf{g}''_{k'}$  by  $-\mathbf{g}''_{k'}$  and replace  $\mathbf{g}''_{k''}$  by  $-\mathbf{g}''_{k''}$  we remain in  $[\mathbf{g}] \cap \mathbb{R}_{\text{diag}}^{S \times K}$  and can only make  $\|\mathbf{g}' - \mathbf{g}''\|_\infty$  decrease. Repeating this process until all the  $\lambda_k$ 's are non-negative, we can assume without loss of generality that

$$\lambda_k \geq 0, \quad \text{whatever } k \in \mathbb{N}_K.$$

This being said, we establish two other simple facts that motivate the structure of the proof. First, in order to find an upper bound for (12), we easily establish (using (8)) that

$$\begin{aligned}
\|\mathbf{g}'_k\|_\infty &= (P(\mathbf{g})_i)^{-1+\frac{1}{K}} \|E_i(P(\mathbf{g}))\|_\infty \\
&\leq \left(\frac{1}{2} \|P(\mathbf{h})\|_\infty\right)^{-1+\frac{1}{K}} \|P(\mathbf{h})\|_\infty \\
&\leq 2 \|P(\mathbf{h})\|_\infty^{\frac{1}{K}}.
\end{aligned} \quad (13)$$

Second, the value  $\lambda_k$  appearing in (12), can be bounded by using bounds on  $\|\mathbf{g}'_k\|_\infty$  and the identity

$$\|\mathbf{g}''_k\|_\infty = \|P(\mathbf{g})\|_\infty^{\frac{1}{K}} = \lambda_k \|\mathbf{g}'_k\|_\infty. \quad (14)$$

Qualitatively, the latter identity indeed guarantees that, as  $\|P(\mathbf{g}) - P(\mathbf{h})\|_\infty$  goes to 0,  $\lambda_k$  goes to 1. Let us now establish this quantitatively.

Recalling that

$$\mathbf{g}' = (P(\mathbf{g})_i)^{-1+\frac{1}{K}} E_i(P(\mathbf{g})),$$

and using (8) again, we obtain

$$\|\mathbf{g}'_k\|_\infty \leq \left( \|P(\mathbf{h})\|_\infty - \frac{1}{2}\|P(\mathbf{h}) - P(\mathbf{g})\|_\infty \right)^{-1+\frac{1}{k}} \|P(\mathbf{g})\|_\infty.$$

We also have (again, using (8))

$$\begin{aligned} \|\mathbf{g}'_k\|_\infty &\geq (P(\mathbf{g})_i)^{-1+\frac{1}{k}} |P(\mathbf{g})_i| \\ &= (P(\mathbf{g})_i)^{\frac{1}{k}} \\ &\geq \left( \|P(\mathbf{h})\|_\infty - \frac{1}{2}\|P(\mathbf{h}) - P(\mathbf{g})\|_\infty \right)^{\frac{1}{k}}. \end{aligned}$$

Plugging the upper bound of  $\|\mathbf{g}'_k\|_\infty$  in (14), using successively (8), the mean value theorem and the hypothesis on the size of  $P(\mathbf{h}) - P(\mathbf{g})$  gives:

$$\begin{aligned} \lambda_k - 1 &= \frac{\|P(\mathbf{g})\|_\infty^{\frac{1}{k}}}{\|\mathbf{g}'_k\|_\infty} - 1 \\ &\geq \|P(\mathbf{g})\|_\infty^{-1+\frac{1}{k}} \left( \|P(\mathbf{h})\|_\infty - \frac{1}{2}\|P(\mathbf{h}) - P(\mathbf{g})\|_\infty \right)^{1-\frac{1}{k}} - 1 \\ &\geq \left( 1 - \frac{\|P(\mathbf{h}) - P(\mathbf{g})\|_\infty}{2\|P(\mathbf{h})\|_\infty} \right)^{1-\frac{1}{k}} - 1 \\ &\geq -\left(1 - \frac{1}{K}\right) \left( 1 - \frac{\|P(\mathbf{h}) - P(\mathbf{g})\|_\infty}{2\|P(\mathbf{h})\|_\infty} \right)^{-\frac{1}{k}} \frac{\|P(\mathbf{h}) - P(\mathbf{g})\|_\infty}{2\|P(\mathbf{h})\|_\infty} \\ &\geq -\left(1 - \frac{1}{4}\right)^{-\frac{1}{k}} \frac{\|P(\mathbf{h}) - P(\mathbf{g})\|_\infty}{2\|P(\mathbf{h})\|_\infty} \\ &\geq -\frac{\|P(\mathbf{h}) - P(\mathbf{g})\|_\infty}{\|P(\mathbf{h})\|_\infty}. \end{aligned}$$

Similarly, plugging the lower bound of  $\|\mathbf{g}'_k\|_\infty$  in (14), we obtain using successively (8), the mean value theorem and the hypothesis on the size of  $P(\mathbf{h}) - P(\mathbf{g})$ :

$$\begin{aligned} \lambda_k - 1 &\leq \|P(\mathbf{g})\|_\infty^{\frac{1}{k}} \left( \|P(\mathbf{h})\|_\infty - \frac{1}{2}\|P(\mathbf{h}) - P(\mathbf{g})\|_\infty \right)^{-\frac{1}{k}} - 1 \\ &\leq \left( 1 - \frac{\|P(\mathbf{h}) - P(\mathbf{g})\|_\infty}{2\|P(\mathbf{h})\|_\infty} \right)^{-\frac{1}{k}} - 1 \\ &\leq \frac{1}{K} \left( 1 - \frac{\|P(\mathbf{h}) - P(\mathbf{g})\|_\infty}{2\|P(\mathbf{h})\|_\infty} \right)^{-1-\frac{1}{k}} \frac{\|P(\mathbf{h}) - P(\mathbf{g})\|_\infty}{2\|P(\mathbf{h})\|_\infty} \\ &\leq \frac{1}{K} \left( 1 - \frac{1}{4} \right)^{-1-\frac{1}{k}} \frac{\|P(\mathbf{h}) - P(\mathbf{g})\|_\infty}{2\|P(\mathbf{h})\|_\infty} \\ &\leq \frac{4^2}{2K3^2} \frac{\|P(\mathbf{h}) - P(\mathbf{g})\|_\infty}{\|P(\mathbf{h})\|_\infty} \\ &\leq \frac{\|P(\mathbf{h}) - P(\mathbf{g})\|_\infty}{\|P(\mathbf{h})\|_\infty}. \end{aligned}$$

Finally, we get

$$|\lambda_k - 1| \leq \frac{\|P(\mathbf{h}) - P(\mathbf{g})\|_\infty}{\|P(\mathbf{h})\|_\infty}. \quad (15)$$

By combining (12), (13) and (15), we obtain

$$\|\mathbf{g}'_k - \mathbf{g}''_k\|_\infty \leq 2 \|P(\mathbf{h})\|_\infty^{-1+\frac{1}{K}} \|P(\mathbf{h}) - P(\mathbf{g})\|_\infty.$$

Combining the latter inequality with (9) and (10) provides

$$d_\infty([\mathbf{h}], [\mathbf{g}]) \leq 7 \|P(\mathbf{h})\|_\infty^{-1+\frac{1}{K}} \|P(\mathbf{h}) - P(\mathbf{g})\|_\infty,$$

and concludes the proof when  $p = q = +\infty$ .

In order to establish the property when  $1 \leq p \leq +\infty$  and  $1 \leq q \leq +\infty$ , we simply use the fact that

$$d_p([\mathbf{h}], [\mathbf{g}]) \leq (KS)^{\frac{1}{p}} d_\infty([\mathbf{h}], [\mathbf{g}])$$

and

$$\|P(\mathbf{h}) - P(\mathbf{g})\|_\infty \leq \|P(\mathbf{h}) - P(\mathbf{g})\|_q.$$

□

The following proposition shows that the upper bound in (7) cannot be improved by a significant factor.

**Proposition 2.** *There exist  $\mathbf{h}$  and  $\mathbf{g} \in \mathbb{R}_*^{S \times K}$  such that  $\|P(\mathbf{g})\|_\infty \leq \|P(\mathbf{h})\|_\infty$ ,  $\|P(\mathbf{g}) - P(\mathbf{h})\|_\infty \leq \frac{1}{2} \|P(\mathbf{h})\|_\infty$  and*

$$7(KS)^{\frac{1}{p}} \|P(\mathbf{h})\|_\infty^{\frac{1}{K}-1} \|P(\mathbf{h}) - P(\mathbf{g})\|_q \leq C_q d_p([\mathbf{h}], [\mathbf{g}]),$$

where

$$C_q = \begin{cases} 28(KS)^{\frac{1}{q}} & , \text{if } q < +\infty, \\ 28 & , \text{if } q = +\infty. \end{cases}$$

*Proof.* In the example, we consider  $\mathbf{h}$  and  $\mathbf{g}$  such that for all  $k \in \mathbb{N}_K$  and all  $i \in \mathbb{N}_S$

$$\mathbf{h}_{k,i} = \begin{cases} 1 & , \text{if } i = 0, \\ 0 & , \text{otherwise,} \end{cases} \quad \text{and} \quad \mathbf{g}_{k,i} = \begin{cases} (\frac{1}{2})^{\frac{1}{K}} & , \text{if } i = 0, \\ \varepsilon_q & , \text{otherwise,} \end{cases}$$

where  $\varepsilon_{+\infty} = (\frac{1}{2})^{\frac{1}{K}}$  and  $\varepsilon_q = \min \left( \left( \frac{1 - (\frac{1}{2})^{\frac{q}{K}}}{S-1} \right)^{\frac{1}{q}}, (\frac{1}{2})^{\frac{1}{K}} \right)$ , if  $q < +\infty$ . We immediately

obtain

$$\|P(\mathbf{h})\|_\infty = 1, \quad \|P(\mathbf{g})\|_\infty = \frac{1}{2} \quad \text{and} \quad \|P(\mathbf{h}) - P(\mathbf{g})\|_\infty = \frac{1}{2}.$$

We also have,

$$d_p([\mathbf{h}], [\mathbf{g}])^p = \|\mathbf{h} - \mathbf{g}\|_p^p \geq K(S-1) \varepsilon_q^p \geq \frac{KS}{2} \varepsilon_q^p.$$

Decomposing the sum necessary to the calculation of the  $l^q$  norm of a tensor according to number of index different from 0 (which corresponds to  $l$  in the sum below), we obtain

$$\begin{aligned}\|P(\mathbf{h}) - P(\mathbf{g})\|_q^q &= \sum_{l=0}^K \binom{l}{K} (S-1)^l \varepsilon_q^{lq} \left(\frac{1}{2}\right)^{\frac{(K-l)q}{K}}, \\ &= \left( \left(\frac{1}{2}\right)^{\frac{q}{K}} + (S-1)\varepsilon_q^q \right)^K \leq 1.\end{aligned}$$

We then easily obtain that

$$\begin{aligned}7\|P(\mathbf{h})\|_\infty^{-1+\frac{1}{K}}(KS)^{\frac{1}{p}}\|P(\mathbf{h}) - P(\mathbf{g})\|_q &\leq 7(KS)^{\frac{1}{p}}, \\ &\leq 7 \frac{d_p([\mathbf{h}], [\mathbf{g}])}{\varepsilon_q} 2^{\frac{1}{p}}.\end{aligned}\tag{16}$$

$$\tag{17}$$

We first calculate a lower bound of  $\varepsilon_q$  when  $\varepsilon_q = \left(\frac{1 - \left(\frac{1}{2}\right)^{\frac{q}{K}}}{S-1}\right)^{\frac{1}{q}}$  (which, in particular, rules out  $q = +\infty$ ). Using the mean value theorem, we obtain

$$1 - \left(\frac{1}{2}\right)^{\frac{q}{K}} \geq \min_{t \in [\frac{1}{2}, 1]} \left(\frac{q}{K} t^{\frac{q}{K}-1}\right) \left(1 - \frac{1}{2}\right).$$

Distinguishing, whether  $q \leq K$  or not, we find after a short calculation that, since  $q \geq 1$ ,

$$1 - \left(\frac{1}{2}\right)^{\frac{q}{K}} \geq \min\left(\frac{1}{2K}, \frac{1}{K} \left(\frac{1}{2}\right)^{\frac{q}{K}}\right) = \frac{1}{K} \min\left(\left(\frac{1}{2}\right)^{\frac{1}{q}}, \left(\frac{1}{2}\right)^{\frac{1}{K}}\right)^q \geq \frac{1}{K2^q}.$$

We then deduce

$$\varepsilon_q \geq \frac{1}{2(KS)^{\frac{1}{q}}}.$$

Of course, when  $\varepsilon_q = \left(\frac{1}{2}\right)^{\frac{1}{K}}$  (which includes  $q = +\infty$ ), we immediately obtain

$$\varepsilon_q \geq \frac{1}{2}.$$

Using this lower bound in (16) leads to the bounds stated in the proposition.  $\square$

When interpreting this proposition one has to remember that, in many practical cases, the product  $KS$  is not very large. Also, the bound is sharp when  $q$  is large.

As stated in the following theorem, we have a more valuable upper bound in the general case.

**Theorem 2. Lipschitz continuity of  $P$**

We have for any  $q \in [1, \infty]$  and any  $\mathbf{h}$  and  $\mathbf{g} \in \mathbb{R}_*^{S \times K}$ ,

$$\|P(\mathbf{h}) - P(\mathbf{g})\|_q \leq S^{\frac{K-1}{q}} K^{1-\frac{1}{q}} \max\left(\|P(\mathbf{h})\|_\infty^{1-\frac{1}{K}}, \|P(\mathbf{g})\|_\infty^{1-\frac{1}{K}}\right) d_q([\mathbf{h}], [\mathbf{g}]).\tag{18}$$

*Proof.* Before starting the proof, we define for any  $k \in \{0, \dots, K\}$

$$P_k(\mathbf{h}, \mathbf{g})_{\mathbf{i}} = \mathbf{g}_{1, \mathbf{i}_1} \dots \mathbf{g}_{k, \mathbf{i}_k} \mathbf{h}_{k+1, \mathbf{i}_{k+1}} \dots \mathbf{h}_{K, \mathbf{i}_K}, \quad \text{for all } \mathbf{h}, \mathbf{g} \in \mathbb{R}^{S \times K} \text{ and all } \mathbf{i} \in \mathbb{N}_S^K.$$

We consider  $\mathbf{g}$  and  $\mathbf{h} \in \mathbb{R}^{S \times K}$ . Let us first assume that  $\|\mathbf{g}\|_\infty \leq \|\mathbf{h}\|_\infty = 1$ . We have for any  $\mathbf{i} \in \mathbb{N}_S^K$ , using this hypothesis and standard inequalities between  $l^p$  norms, when  $q < +\infty$

$$\begin{aligned} |P(\mathbf{g})_{\mathbf{i}} - P(\mathbf{h})_{\mathbf{i}}|^q &= \left| \sum_{k=0}^{K-1} (P_{k+1}(\mathbf{h}, \mathbf{g})_{\mathbf{i}} - P_k(\mathbf{h}, \mathbf{g})_{\mathbf{i}}) \right|^q \\ &\leq K^{q-1} \sum_{k=0}^{K-1} |P_{k+1}(\mathbf{h}, \mathbf{g})_{\mathbf{i}} - P_k(\mathbf{h}, \mathbf{g})_{\mathbf{i}}|^q \\ &\leq K^{q-1} \sum_{k=0}^{K-1} |\mathbf{g}_{k+1, \mathbf{i}_{k+1}} - \mathbf{h}_{k+1, \mathbf{i}_{k+1}}|^q \end{aligned}$$

The same calculation when  $q = +\infty$  leads to

$$|P(\mathbf{g})_{\mathbf{i}} - P(\mathbf{h})_{\mathbf{i}}| \leq K \max_{k=1..K} |\mathbf{g}_{k, \mathbf{i}_k} - \mathbf{h}_{k, \mathbf{i}_k}|.$$

Therefore, we have when  $q < +\infty$

$$\begin{aligned} \|P(\mathbf{h}) - P(\mathbf{g})\|_q^q &= \sum_{\mathbf{i} \in \mathbb{N}_S^K} |P(\mathbf{h})_{\mathbf{i}} - P(\mathbf{g})_{\mathbf{i}}|^q \\ &\leq K^{q-1} \sum_{k=1}^K \sum_{\mathbf{i} \in \mathbb{N}_S^K} |\mathbf{g}_{k, \mathbf{i}_k} - \mathbf{h}_{k, \mathbf{i}_k}|^q \\ &= K^{q-1} \sum_{k=1}^K S^{K-1} \|\mathbf{g}_k - \mathbf{h}_k\|_q^q \\ &= K^{q-1} S^{K-1} \|\mathbf{g} - \mathbf{h}\|_q^q \end{aligned}$$

and therefore

$$\|P(\mathbf{h}) - P(\mathbf{g})\|_q \leq K^{1-\frac{1}{q}} S^{\frac{K-1}{q}} \|\mathbf{g} - \mathbf{h}\|_q.$$

Again, a similar calculus for  $q = +\infty$  leads to

$$\|P(\mathbf{h}) - P(\mathbf{g})\|_{+\infty} \leq K \|\mathbf{g} - \mathbf{h}\|_{+\infty}.$$

Remember that the two last inequalities hold for  $\mathbf{g}$  and  $\mathbf{h} \in \mathbb{R}^{S \times K}$  such that  $\|\mathbf{g}\|_\infty \leq \|\mathbf{h}\|_\infty = 1$ .

Let us now consider any  $\mathbf{g}'$  and  $\mathbf{h}' \in \mathbb{R}^{S \times K}$  and any  $\mathbf{g} \in \mathbb{R}_{\text{diag}}^{S \times K} \cap [\mathbf{g}']$  and  $\mathbf{h} \in \mathbb{R}_{\text{diag}}^{S \times K} \cap [\mathbf{h}']$ .

We denote  $\delta = \max(\|\mathbf{g}\|_{+\infty}, \|\mathbf{h}\|_{+\infty})$ . Notice first that  $\|\mathbf{g}\|_{+\infty} = \|P(\mathbf{g}')\|_{+\infty}^{\frac{1}{K}}$  and  $\|\mathbf{h}\|_{+\infty} = \|P(\mathbf{h}')\|_{+\infty}^{\frac{1}{K}}$ . Therefore

$$\delta = \max(\|P(\mathbf{g}')\|_{+\infty}, \|P(\mathbf{h}')\|_{+\infty})^{\frac{1}{K}}. \quad (19)$$

We can apply the above inequality to  $\frac{\mathbf{h}}{\delta}$  and  $\frac{\mathbf{g}}{\delta}$  (we might need to switch  $\mathbf{h}$  and  $\mathbf{g}$  but it does not change the final inequality) and obtain when  $q < +\infty$

$$\left\| P\left(\frac{\mathbf{h}}{\delta}\right) - P\left(\frac{\mathbf{g}}{\delta}\right) \right\|_q \leq K^{1-\frac{1}{q}} S^{\frac{K-1}{q}} \left\| \frac{\mathbf{g}}{\delta} - \frac{\mathbf{h}}{\delta} \right\|_q.$$

This leads to

$$\|P(\mathbf{h}) - P(\mathbf{g})\|_q \leq K^{1-\frac{1}{q}} S^{\frac{K-1}{q}} \delta^{K-1} \|\mathbf{g} - \mathbf{h}\|_q.$$

Similarly, when  $q = +\infty$ , we obtain

$$\|P(\mathbf{h}) - P(\mathbf{g})\|_{+\infty} \leq K\delta^{K-1} \|\mathbf{g} - \mathbf{h}\|_{+\infty}.$$

The fact that these two last inequalities hold for any  $\mathbf{g} \in \mathbb{R}_{\text{diag}}^{S \times K} \cap [\mathbf{g}']$  and any  $\mathbf{h} \in \mathbb{R}_{\text{diag}}^{S \times K} \cap [\mathbf{h}']$ , together with (19), leads to the statement provided in Theorem 2.  $\square$

Notice that, considering  $\mathbf{h}$  and  $\mathbf{g} \in \mathbb{R}^{S \times K}$  such that  $\mathbf{h}_{k,i} = 1$  and  $\mathbf{g}_{k,i} = \varepsilon$ , for all  $k \in \mathbb{N}_K$  and  $i \in \mathbb{N}_S$  and for a  $0 < \varepsilon \ll 1$ , we easily calculate

$$S^{\frac{K-1}{q}} K^{1-\frac{1}{q}} \max \left( \|P(\mathbf{h})\|_{\infty}^{1-\frac{1}{K}}, \|P(\mathbf{g})\|_{\infty}^{1-\frac{1}{K}} \right) d_q([\mathbf{h}], [\mathbf{g}]) \leq K \|P(\mathbf{h}) - P(\mathbf{g})\|_q.$$

As a consequence, the upper bound in Theorem 2 is tight up to at most a factor  $K$ . This factor is not significant since, in the present context,  $K$  is not intended to become significantly large.

## 4 The lifting Principle

The following proposition is clear (it can be shown by induction on  $K$ ):

**Proposition 3.** *The entries of the matrix*

$$M_1(\mathbf{h}_1)M_2(\mathbf{h}_2) \dots M_K(\mathbf{h}_K)$$

are multivariate polynomials whose variables are the entries of  $\mathbf{h} \in \mathbb{R}^{S \times K}$ . Moreover, every entry is the sum of monomials of degree  $K$ . Each monomial is a constant times  $\mathbf{h}_{1,i_1} \dots \mathbf{h}_{K,i_K}$ , for some  $\mathbf{i} \in \mathbb{N}_S^K$ .

Notice that any monomial  $\mathbf{h}_{1,i_1} \dots \mathbf{h}_{K,i_K}$  is the entry  $P(\mathbf{h})_{\mathbf{i}}$  in the tensor  $P(\mathbf{h})$ . Therefore every polynomial in the previous proposition takes the form  $\sum_{\mathbf{i} \in \mathbb{N}_S^K} c_{\mathbf{i}} P(\mathbf{h})_{\mathbf{i}}$  for some constants  $(c_{\mathbf{i}})_{\mathbf{i} \in \mathbb{N}_S^K}$  independent of  $\mathbf{h}$ . In words, every entry of the matrix  $M_1(\mathbf{h}_1)M_2(\mathbf{h}_2) \dots M_K(\mathbf{h}_K)$  is obtained by applying a linear form to  $P(\mathbf{h})$ . Moreover, the polynomial coefficients defining the linear form are uniquely determined by the linear maps  $M_1, \dots, M_K$ . This leads to the following statement.

**Corollary 1.** *Let  $M_k$ ,  $k \in \mathbb{N}_K$  be as in (1). The map*

$$(\mathbf{h}_1, \dots, \mathbf{h}_K) \mapsto M_1(\mathbf{h}_1)M_2(\mathbf{h}_2) \dots M_K(\mathbf{h}_K),$$

uniquely determines a linear map

$$\mathcal{A} : \mathbb{R}^{S^K} \longrightarrow \mathbb{R}^{m \times n},$$

such that for all  $\mathbf{h} \in \mathbb{R}^{S \times K}$

$$M_1(\mathbf{h}_1)M_2(\mathbf{h}_2) \dots M_K(\mathbf{h}_K) = \mathcal{A}P(\mathbf{h}). \quad (20)$$

When  $K = 1$ , we simply have  $\mathcal{A} = M_1$ . When  $K \geq 2$ , it may be difficult to provide a closed form expression for the operator  $\mathcal{A}$ . We can however determine simple properties of  $\mathcal{A}$ . In most reasonable cases,  $\mathcal{A}$  is sparse. If the operators  $M_k$  simply embed the values of  $h$  in a matrix, the matrix representing  $\mathcal{A}$  only contains zeros and ones. Also, since the



operators  $M_k$  are known, we can compute  $\mathcal{A}P(\mathbf{h})$ , whatever  $\mathbf{h} \in \mathbb{R}^{S \times K}$ , using (20). Said differently, we can compute  $\mathcal{A}$  for any rank 1 entry. Therefore, since  $\mathcal{A}$  is linear, we can compute  $\mathcal{A}T$  for any low rank tensor  $T$ . If the dimensions of the problem permit, one can manipulate  $\mathcal{A}$  in a basis of  $\mathbb{R}^{S^K}$ .

We also have a way to compute  $\text{rk}(\mathcal{A})$ , when  $mn$  is not too large, using the following proposition.

**Proposition 4.** *If we consider  $R$  independent random collections of vectors  $\mathbf{h}^r$ , with  $r = 1 \dots R$ , according to the normal distribution in  $\mathbb{R}^{S \times K}$ , we have (with probability 1)*

$$\dim(\text{Span}((\mathcal{A}P(\mathbf{h}^r))_{r=1..R})) = \begin{cases} R & , \text{ if } R \leq \text{rk}(\mathcal{A}) \\ \text{rk}(\mathcal{A}) & , \text{ otherwise.} \end{cases} \quad (21)$$

*Proof.* The span of the Segre variety  $P(\mathbb{R}^{S \times K})$  is the full ambient space  $\mathbb{R}^{S^K}$ , so there exists sets of  $R \leq S^K$  points on it that are linearly independent. The set of  $R$ -tuples of points on  $P(\mathbb{R}^{S \times K})$  that fail to be linearly independent is a proper subvariety of the variety of sets of  $R$ -tuples of points on  $P(\mathbb{R}^{S \times K})$  because being a linearly independent set of points is an open condition and there exists sets of points that are linearly independent. Therefore  $R \leq S^K$  independent and randomly chosen points according to a continuous distribution on  $P(\mathbb{R}^{S \times K})$  will be linearly independent.

The intersection  $P(\mathbb{R}^{S \times K}) \cap \text{Ker}(\mathcal{A})$  is a proper subvariety of  $P(\mathbb{R}^{S \times K})$ , so with probability one,  $R \leq S^K$  independent randomly chosen points according to a continuous distribution will not intersect it and be linearly independent. This is indeed the intersection of two non-empty open conditions. Therefore, all spans of subsets of the points will intersect  $\text{Ker}(\mathcal{A})$  transversely (in particular, the span of fewer than  $\text{rk}(\mathcal{A})$  points will not intersect it). Thus there image under  $\mathcal{A}$  will have dimension as large as possible. The same argument works if  $R > S^K$ .  $\square$

Using Corollary 1, when (2) has a minimizer, we rewrite the problem (2) in the form

$$\mathbf{h}^* \in \underset{L \in \mathbb{N}, \mathbf{h} \in \mathcal{M}^L}{\text{argmin}} \|\mathcal{A}P(\mathbf{h}) - X\|^2. \quad (22)$$

We now decompose this problem into two sub-problems: A least-squares problem

$$T^* \in \underset{T \in \mathbb{R}^{S^K}}{\text{argmin}} \|\mathcal{A}T - X\|^2 \quad (23)$$

and a non-convex problem

$$\mathbf{h}^{/*} \in \underset{L \in \mathbb{N}, \mathbf{h} \in \mathcal{M}^L}{\text{argmin}} \|\mathcal{A}(P(\mathbf{h}) - T^*)\|^2. \quad (24)$$

**Proposition 5.** *When (2) has a minimizer, for any  $X$ ,  $\mathcal{A}$ :*

1. *Let  $\mathbf{h}^*$  be a solution of (22). Then, for any solution  $T^*$  of (23),  $\mathbf{h}^*$  also minimizes (24).*
2. *Let  $T^*$  be a solution of (23) and  $\mathbf{h}^{/*}$  a solution of (24). Then,  $\mathbf{h}^{/*}$  also minimizes (22).*

*Proof.* The proof relies on the fact that for any  $T^* \in \underset{T \in \mathbb{R}^{S^K}}{\text{argmin}} \|\mathcal{A}T - X\|^2$ , we have

$$\mathcal{A}^t(\mathcal{A}T^* - X) = 0,$$

where  $\mathcal{A}^t : \mathbb{R}^{n \times m} \rightarrow \mathbb{R}^{S^K}$  is the adjoint linear map. This implies that for any

$$T^* \in \underset{T \in \mathbb{R}^{S^K}}{\text{argmin}} \|\mathcal{A}T - X\|^2,$$

any  $L \in \mathbb{N}$  and any  $\mathbf{h} \in \mathcal{M}^L$

$$\begin{aligned}\|\mathcal{A}P(\mathbf{h}) - X\|^2 &= \|\mathcal{A}(P(\mathbf{h}) - T^*) + (\mathcal{A}T^* - X)\|^2, \\ &= \|\mathcal{A}(P(\mathbf{h}) - T^*)\|^2 + \|\mathcal{A}T^* - X\|^2 + 2\langle \mathcal{A}(P(\mathbf{h}) - T^*), \mathcal{A}T^* - X \rangle, \\ &= \|\mathcal{A}(P(\mathbf{h}) - T^*)\|^2 + \|\mathcal{A}T^* - X\|^2.\end{aligned}$$

In words,  $\|\mathcal{A}P(\mathbf{h}) - X\|^2$  and  $\|\mathcal{A}(P(\mathbf{h}) - T^*)\|^2$  only differ by an additive constant. Moreover, since the value of the objective function  $\|\mathcal{A}T^* - X\|^2$  is independent of the particular minimizer  $T^*$  we are considering, this additive constant is independent of  $T^*$ . As a consequence, a minimizer of  $\|\mathcal{A}P(\mathbf{h}) - X\|^2$  also minimizes  $\|\mathcal{A}(P(\mathbf{h}) - T^*)\|^2$  and vice versa.  $\square$

From now on, because of the equivalence between solutions of (24) and (22), we stop using the notation  $\mathbf{h}^*$  and write  $\mathbf{h}^* \in \operatorname{argmin}_{L \in \mathbb{N}, \mathbf{h} \in \mathcal{M}^L} \|\mathcal{A}(P(\mathbf{h}) - T^*)\|^2$ .

## 5 Identifiability in the noise free case

Throughout this section, we assume that  $X$  is such that there exists  $\bar{L}$  and  $\bar{\mathbf{h}} \in \mathcal{M}^{\bar{L}}$  such that

$$X = M_1(\bar{\mathbf{h}}_1) \dots M_K(\bar{\mathbf{h}}_K). \quad (25)$$

Under this assumption,  $X = \mathcal{A}P(\bar{\mathbf{h}})$ , so

$$P(\bar{\mathbf{h}}) \in \operatorname{argmin}_{T \in \mathbb{R}^{s_K}} \|\mathcal{A}T - X\|^2.$$

Moreover, we trivially have  $P(\bar{\mathbf{h}}) \in \Sigma_1$  and therefore  $\bar{\mathbf{h}}$  minimizes (24), (2) and (22). As a consequence, (2) has a minimizer.

We ask whether there exist guarantees that the resolution of (2) allows one to recover  $\bar{\mathbf{h}}$  (up to the usual uncertainties).

In this regard, for any  $\mathbf{h} \in [\bar{\mathbf{h}}]$ , we have  $P(\mathbf{h}) = P(\bar{\mathbf{h}})$  and therefore  $\mathcal{A}P(\mathbf{h}) = \mathcal{A}P(\bar{\mathbf{h}}) = X$ . Thus unless we make further assumptions on  $\bar{\mathbf{h}}$ , we cannot expect to distinguish any particular element of  $[\bar{\mathbf{h}}]$  using only  $X$ . In other words, recovering  $[\bar{\mathbf{h}}]$  is the best we can hope for.

### Definition 2. Identifiability

We say that  $[\bar{\mathbf{h}}]$  is identifiable if the elements of  $[\bar{\mathbf{h}}]$  are the only solutions of (2).

We say that  $\mathcal{M}$  is identifiable if for every  $L \in \mathbb{N}$  and every  $\bar{\mathbf{h}} \in \mathcal{M}^L$ ,  $[\bar{\mathbf{h}}]$  is identifiable.

### Proposition 6. Characterization of the global minimizers

For any  $L^* \in \mathbb{N}$  and any  $\mathbf{h}^* \in \mathcal{M}^{L^*}$ ,  $(L^*, \mathbf{h}^*) \in \operatorname{argmin}_{L \in \mathbb{N}, \mathbf{h} \in \mathcal{M}} \|\mathcal{A}P(\mathbf{h}) - X\|^2$  if and only if

$$P(\mathbf{h}^*) \in P(\bar{\mathbf{h}}) + \operatorname{Ker}(\mathcal{A}).$$

*Proof.* Write  $\bar{T} = P(\bar{\mathbf{h}})$  and let  $L^*$  and  $\mathbf{h}^*$  be a minimizer of (2). Proposition 5 and the fact that  $\bar{T}$  minimizes (23) implies that  $(L^*, \mathbf{h}^*) \in \operatorname{argmin}_{L \in \mathbb{N}, \mathbf{h} \in \mathcal{M}^L} \|\mathcal{A}(P(\mathbf{h}) - \bar{T})\|^2$ . As a consequence,

$$\|\mathcal{A}(P(\mathbf{h}^*) - \bar{T})\|^2 = 0$$

and

$$P(\mathbf{h}^*) \in \bar{T} + \operatorname{Ker}(\mathcal{A}),$$

proving the first implication.

Conversely, let  $L^* \in \mathbb{N}$  and  $\mathbf{h}^* \in \mathcal{M}^{L^*}$  be such that  $P(\mathbf{h}^*) \in \bar{T} + \text{Ker}(\mathcal{A})$ , then

$$\|\mathcal{A}(P(\mathbf{h}^*) - \bar{T})\|^2 = 0 = \min_{L \in \mathbb{N}, \mathbf{h} \in \mathcal{M}^L} \|\mathcal{A}(P(\mathbf{h}) - \bar{T})\|^2.$$

As a consequence,  $(L^*, \mathbf{h}^*) \in \text{argmin}_{L \in \mathbb{N}, \mathbf{h} \in \mathcal{M}^L} \|\mathcal{A}(P(\mathbf{h}) - \bar{T})\|^2$  and, using Proposition 5,  $\mathbf{h}^*$  is a minimizer of (2).  $\square$

In order to state the following proposition, we define for any  $L$  and  $L' \in \mathbb{N}$

$$P(\mathcal{M}^L) - P(\mathcal{M}^{L'}) := \left\{ P(\mathbf{h}) - P(\mathbf{g}) \mid \mathbf{h} \in \mathcal{M}^L \text{ and } \mathbf{g} \in \mathcal{M}^{L'} \right\} \subset \mathbb{R}^{S^K}.$$

**Proposition 7. Necessary and sufficient conditions of identifiability**

1. For any  $\bar{L}$  and  $\bar{\mathbf{h}} \in \mathcal{M}^{\bar{L}}$ :  $[\bar{\mathbf{h}}]$  is identifiable if and only if for any  $L \in \mathbb{N}$

$$(P(\bar{\mathbf{h}}) + \text{Ker}(\mathcal{A})) \cap P(\mathcal{M}^L) \subset \{P(\bar{\mathbf{h}})\}.$$

2.  $\mathcal{M}$  is identifiable if and only if for any  $L$  and  $L' \in \mathbb{N}$

$$\text{Ker}(\mathcal{A}) \cap (P(\mathcal{M}^L) - P(\mathcal{M}^{L'})) = \{0\}. \quad (26)$$

*Proof.* • Proof of the first statement of Proposition 7:

We first assume that  $[\bar{\mathbf{h}}]$  is identifiable. We consider  $L^*$  and  $\mathbf{h}^*$  such that there is  $L^*$  such that  $P(\mathbf{h}^*) \in (P(\bar{\mathbf{h}}) + \text{Ker}(\mathcal{A})) \cap P(\mathcal{M}^{L^*})$ . We know from Proposition 6 that  $\mathbf{h}^* \in \text{argmin}_{L \in \mathbb{N}, \mathbf{h} \in \mathcal{M}^L} \|\mathcal{A}P(\mathbf{h}) - X\|^2$ . Using that  $[\bar{\mathbf{h}}]$  is identifiable,  $[\mathbf{h}^*] = [\bar{\mathbf{h}}]$  and, from Standard Fact 1 (at the beginning of Section 3), we get  $P(\mathbf{h}^*) = P(\bar{\mathbf{h}})$ . Finally, we can conclude, that if  $[\bar{\mathbf{h}}]$  is identifiable we have  $(P(\bar{\mathbf{h}}) + \text{Ker}(\mathcal{A})) \cap P(\mathcal{M}) \subset \{P(\bar{\mathbf{h}})\}$ . Let us assume now that for all  $L \in \mathbb{N}$ ,  $(P(\bar{\mathbf{h}}) + \text{Ker}(\mathcal{A})) \cap P(\mathcal{M}^L) \subset \{P(\bar{\mathbf{h}})\}$  and consider

$$(L^*, \mathbf{h}^*) \in \text{argmin}_{L \in \mathbb{N}, \mathbf{h} \in \mathcal{M}} \|\mathcal{A}P(\mathbf{h}) - X\|^2.$$

Using Proposition 6, we know that  $P(\mathbf{h}^*) \in (P(\bar{\mathbf{h}}) + \text{Ker}(\mathcal{A})) \cap P(\mathcal{M}^{L^*})$ . Using the hypothesis, we have  $P(\mathbf{h}^*) = P(\bar{\mathbf{h}})$  and using Standard Fact 1, we finally conclude that  $[\mathbf{h}^*] = [\bar{\mathbf{h}}]$ . This completes the proof of the first statement.

- Proof of the second statement of Proposition 7:

If we assume that there is  $L$  and  $L' \in \mathbb{N}$  such that  $\text{Ker}(\mathcal{A}) \cap (P(\mathcal{M}^L) - P(\mathcal{M}^{L'})) \neq \{0\}$  then there exist  $\mathbf{h} \in \mathcal{M}^L$  and  $\bar{\mathbf{h}} \in \mathcal{M}^{L'}$  such that  $P(\mathbf{h}) \neq P(\bar{\mathbf{h}})$  and  $P(\mathbf{h}) - P(\bar{\mathbf{h}}) \in \text{Ker}(\mathcal{A})$ . Using the first statement of the proposition, we obtain that  $[\bar{\mathbf{h}}]$  is not identifiable. As a conclusion,  $\mathcal{M}$  is not identifiable.

Conversely, if we assume that there exists  $L'$  and some non-identifiable  $\bar{\mathbf{h}} \in \mathcal{M}^{L'}$ . Using the first statement of the proposition, we know that there exists  $L \in \mathbb{N}$  and  $\mathbf{h} \in \mathcal{M}^L$  such that  $P(\mathbf{h}) \neq P(\bar{\mathbf{h}})$  and  $P(\mathbf{h}) - P(\bar{\mathbf{h}}) \in \text{Ker}(\mathcal{A})$ . Therefore  $\text{Ker}(\mathcal{A}) \cap (P(\mathcal{M}) - P(\mathcal{M}^{L'})) \neq \{0\}$ .  $\square$

In the context of the usual compressed sensing (i.e., when  $K = 1$ ,  $\mathcal{M}$  contains  $L$ -sparse signals,  $\mathcal{A}$  is a rectangular matrix with full row rank and  $X$  is a vector), the proposition is already stated in Lemma 3.1 of [13].

In reasonably small cases and when  $P(\mathcal{M})$  is algebraic, one can use tools from numerical algebraic geometry such as those described in [23, 24] to check whether the condition

(26) holds or not. The drawback of Proposition 7 is that, given a factorization model described by  $\mathcal{A}$ , the condition (26) might be difficult to verify.

We therefore establish simpler conditions related to the identifiability of  $\mathcal{M}$ . First we establish a condition such that for almost every  $\mathcal{A}$  satisfying it,  $\mathcal{M}$  is identifiable. The main benefit of this condition is that its constituents can be computed in many practical situations.

Before that, we recall a few facts of algebraic geometry, for  $X, Y \subset \mathbb{R}^N$ , the *join* of  $X$  and  $Y$  (see, e.g., [22, Ex. 8.1]) is

$$J(X, Y) := \overline{\{sx + ty \mid x \in X, y \in Y, s, t \in \mathbb{R}\}}^{\text{Zar}}.$$

If for all  $L \in \mathbb{N}$ ,  $\mathcal{M}^L$  is Zariski closed and invariant under rescaling (e.g., if they are all linear spaces), then  $P(\mathcal{M}^L) - P(\mathcal{M}^{L'})$  is a Zariski open subset of  $J(P(\mathcal{M}^L), P(\mathcal{M}^{L'}))$ . In general, it is contained in this join.

Recall the following fact (\*): for complex algebraic varieties  $X, Y \subset \mathbb{C}^N$ , any component  $Z$  of  $X \cap Y$  has  $\dim(Z) \geq \dim(X) + \dim(Y) - N$ , and equality holds generically (we make “generically” precise in our context below). Moreover, if  $X, Y$  are invariant under rescaling, since  $0 \in X \cap Y$ , we have  $X \cap Y \neq \emptyset$ . (See, e.g., [44, §I.6.2].)

This intersection result indicates that if there exists  $L, L'$  such that

$$\text{rk}(\mathcal{A}) < \dim\left(P(\mathcal{M}^L) - P(\mathcal{M}^{L'})\right)$$

we expect to have non-identifiability; and if the rank is larger, for all pair  $L, L'$ , we expect identifiability.

It is straightforward to make the identifiability assertion precise:

**Theorem 3. Almost surely sufficient condition for Identifiability**

*For almost every  $\mathcal{A}$  such, for all  $L, L'$ , that  $\text{rk}(\mathcal{A}) \geq \dim\left(J(P(\mathcal{M}^L), P(\mathcal{M}^{L'}))\right)$ ,  $\mathcal{M}$  is identifiable.*

*Proof.* We first make the “equality holds generically” statement precise in our context. Fix any variety  $X$  and assume  $Y$  is a linear space, say of dimension  $y$ . Let  $G(y, \mathbb{C}^N)$  denote the Grassmannian of  $y$ -planes through the origin in  $\mathbb{C}^N$ . The Grassmannian is both a smooth manifold and an algebraic variety. We can interpret “equality holds generically” in this context as saying for a Zariski open subset of  $G(y, \mathbb{C}^N)$ , equality will hold. In our situation, if we fix  $\text{rk}(\mathcal{A})$  and allow  $\ker(\mathcal{A})$  to vary as a point in the Grassmannian, with probability one, it will intersect  $J(P(\mathcal{M}^L), P(\mathcal{M}^{L'}))$  only in the origin, and this assertion is also true over  $\mathbb{R}$  because complex numbers are only needed to assure existence of intersections, not non-existence.  $\square$

Since  $\dim\left(J(P(\mathcal{M}^L), P(\mathcal{M}^{L'}))\right) < \dim(P(\mathcal{M}^L)) + \dim(P(\mathcal{M}^{L'})) + 1$ , if  $D_{\max}$  is the maximum dimension of  $P(\mathcal{M}^L)$  over all  $L$ , one has the same conclusion if  $\text{rk}(\mathcal{A}) \geq 2D_{\max} + 1$ .

When  $K = 1$ , we illustrate this result by interpreting it in the context of compressive sensing, where  $\mathbf{h}$  is a vector,  $X$  is a vector,  $\mathcal{A}$  is a rectangular sampling matrix of full row rank and  $\text{Ker}(\mathcal{A})$  is large. The statement analogous to Theorem 3 in the compressive sensing framework takes the form: “For almost every sampling matrix, any  $L$  sparse signal  $\mathbf{h}$  can be recovered from  $\mathcal{A}\mathbf{h}$  as soon as  $2L \leq \text{rk}(\mathcal{A})$ .” Moreover, the constituent of the  $\ell^0$  minimization model used to recover the signal are also the constituents of (22). Again, the main novelty is to extend this result to the identifiability of the factors of a deep matrix products.

In order to establish a necessary condition for identifiability, first note that if we extend  $P(\mathcal{M}^L) - P(\mathcal{M}^{L'})$  to be scale invariant, this will not effect whether or not it intersects  $\ker(\mathcal{A})$  outside of the origin. We immediately conclude that in the complex setting where  $\mathcal{M}^L, \mathcal{M}^{L'}$  are both Zariski closed, that  $\mathcal{M}$  is non-identifiable whenever  $\text{rk}(\mathcal{A}) < \dim\left(P(\mathcal{M}^L) - P(\mathcal{M}^{L'})\right)$ . This indicates that we should always expect non-identifiability whenever  $\text{rk}(\mathcal{A}) < \dim\left(P(\mathcal{M}^L) - P(\mathcal{M}^{L'})\right)$  but is not adequate to prove it because real algebraic varieties need not satisfy (\*). However it is true for real linear spaces, so we immediately conclude the following weak result:

**Theorem 4. Necessary condition for Identifiability**

*Let  $C(P(\mathcal{M}^L) - P(\mathcal{M}^{L'}))$  be the set of all points on all lines through the origin intersecting  $P(\mathcal{M}^L) - P(\mathcal{M}^{L'})$ , and let  $q$  be the maximal dimension of a linear space on  $C(P(\mathcal{M}^L) - P(\mathcal{M}^{L'}))$ . Then if  $q > \text{rk}(\mathcal{A})$ ,  $\mathcal{M}$  is not identifiable. In particular when the  $\mathcal{M}^L$ s contain linear space and if we let  $S'$  be the largest dimension of these vector space, if  $2S' > \text{rk}(\mathcal{A})$ , then  $\mathcal{M}$  is not identifiable.*

## 6 Stable recovery in the noisy case

In this section, we significantly complexify the problem by considering a noisy situation and by considering an inaccurate optimization. More precisely, we assume that we know a sequence  $\overline{\mathcal{M}} = (\mathcal{M}^L)_{L \in \mathbb{N}}$  made of models  $\mathcal{M}^L \subset \mathbb{R}^{S \times K}$ , for  $L \in \mathbb{N}$ . Of course, if the true model  $\overline{\mathcal{M}}$  is known, we simply take  $\mathcal{M}^L = \overline{\mathcal{M}}$  for all  $L \in \mathbb{N}$ . This should simplify the construction of the solver without modifying the statements below.

In this section, we assume that there exists  $\overline{L} \in \mathbb{N}$  and  $\overline{\mathbf{h}} \in \mathcal{M}^{\overline{L}}$ . Our goal is to estimate  $\overline{\mathbf{h}}$  from the noisy matrix product

$$X = M_1(\overline{\mathbf{h}}_1) \dots M_K(\overline{\mathbf{h}}_K) + e$$

for some error  $e$  satisfying

$$\|e\| \leq \delta, \tag{27}$$

for  $\delta > 0$ .

This can equivalently be written

$$X = \mathcal{A}P(\overline{\mathbf{h}}) + e.$$

We consider an inexact minimization as considered in [7]. More precisely, we assume that we have a way to find  $L^*$  and  $\mathbf{h}^* \in \mathcal{M}^{L^*}$  such that

$$\|\mathcal{A}P(\mathbf{h}^*) - X\| \leq \eta, \tag{28}$$

for some parameter  $\eta > 0$ . We will see in Theorem 5 that smaller values of  $\eta$  improve the estimate accuracy.

Typically, this can be achieved by approximatively solving the optimization problem:

$$\operatorname{argmin}_{L \in \mathbb{N}, \mathbf{h} \in \mathcal{M}^L} \|\mathcal{A}P(\mathbf{h}) - X\|^2. \tag{29}$$

The hypothesis  $\mathbf{h}^* \in \mathcal{M}^{L^*}$  might give the false impression that the results described below are restricted to heuristics imposing the constraint. It is not. Indeed, even if  $\mathbf{h}' \in \mathbb{R}^{S \times K}$  is such that

$$\|\mathcal{A}P(\mathbf{h}') - X\| \leq \eta_1,$$

it is possible to find  $L^*$  and  $\mathbf{h}^* \in \mathcal{M}^{L^*}$  such that  $d_2([\mathbf{h}'], [\mathbf{h}^*]) \leq \eta_2$ , where  $\eta_2$  quantifies how much  $\mathbf{h}'$  deviates from the model. Using Theorem 2, we immediately have

$$\|\mathcal{A}P(\mathbf{h}^*) - X\| \leq \eta_1 + \sigma_{\max} S^{\frac{K-1}{2}} \sqrt{K} \max \left( \|P(\mathbf{h}')\|_{\infty}^{1-\frac{1}{K}}, \|P(\mathbf{g}^*)\|_{\infty}^{1-\frac{1}{K}} \right) \eta_2,$$

where  $\sigma_{\max}$  is the spectral radius of  $\mathcal{A}$ . The analysis below can be applied using this bound in place of  $\eta$ .

However, we need not solve this optimization problem (which is in general difficult or impossible to solve). What matters is just to satisfy (28) for a small value of  $\eta$ . Notice that, if we efficiently solve the above problem, we can reach  $\eta \leq \delta$ . Getting a value of  $\eta$  whose order of magnitude is comparable to  $\delta$  is what one should aim at.

Also, notice that dealing with inexact minimization is important since it permits one to avoid the use of the coercivity hypothesis (which might be difficult to certify in practical applications). More generally, it permits one to consider the case when (29) does not have a solution. Moreover, the above inequality always holds when  $\mathbf{h}^*$  is the solution returned by a heuristic algorithm and  $\eta = \|\mathcal{A}P(\mathbf{h}^*) - X\|^2$ . Algorithms which do not come with a guarantee sometimes manage to reach small objective function values. In those cases, the analysis we conduct in this section permits one to conclude the stable recovery of  $\bar{\mathbf{h}}$  despite the lack of a guarantee of the algorithm. Finally, considering an inaccurate minimization permits one to obtain guarantees for algorithms that, instead of minimizing (29), minimize an objective function which approximates the one in (29). The goal of the approximation would be to consider an objective function that better lends itself to optimization or that is accessible using the available data.

A necessary and sufficient condition for the identifiability of  $\mathcal{M}$  is stated in Proposition 7. The condition is on the way  $\text{Ker}(\mathcal{A})$  and  $P(\mathcal{M}^L) - P(\mathcal{M}^{L'})$  intersect. In order to guarantee the stable recovery of the elements of  $\mathcal{M}$ , we need a stronger condition on the geometry of this intersection to hold for every  $L$  and  $L' \in \mathbb{N}$ . This condition is provided in the next definition.

**Definition 3. Deep-Null Space Property**

Let  $\gamma > 0$ , we say that  $\text{Ker}(\mathcal{A})$  satisfies the deep-Null Space Property (deep-NSP) with respect to the model collection  $\mathcal{M}$  with constant  $\gamma$  if there exists  $\varepsilon > 0$  such that for any  $L$  and  $L' \in \mathbb{N}$ , any  $T \in P(\mathcal{M}^L) - P(\mathcal{M}^{L'})$  satisfying  $\|\mathcal{A}T\| \leq \varepsilon$  and any  $T' \in \text{Ker}(\mathcal{A})$ , we have

$$\|T\| \leq \gamma \|T - T'\|. \quad (30)$$

The deep-NSP implies that, for  $T \in P(\mathcal{M}^L) - P(\mathcal{M}^{L'})$  close to  $\text{Ker}(\mathcal{A})$  in the sense that  $\|\mathcal{A}T\| \leq \varepsilon$  we must have, by decomposing  $T = T' + T''$ , with  $T' \in \text{Ker}(\mathcal{A})$  and  $T''$  in its orthogonal complement

$$\|T\| \leq \gamma \|T - T'\| = \gamma \|T''\| \leq \frac{\gamma}{\sigma_{\min}} \|\mathcal{A}T''\| \leq \frac{\gamma}{\sigma_{\min}} \varepsilon,$$

where  $\sigma_{\min}$  is the smallest non-zero singular value of  $\mathcal{A}$ . In words,  $\|T\|$  must be small. We can conclude that under the deep-NSP,  $P(\mathcal{M}^L) - P(\mathcal{M}^{L'})$  and  $\{T \in \mathbb{R}^{S^K} \mid \|\mathcal{A}T\| \leq \varepsilon\}$  only intersect in the vicinity of 0.

Additionally, (30) implies that in the vicinity of 0,  $\text{Ker}(\mathcal{A})$  and  $P(\mathcal{M}^L) - P(\mathcal{M}^{L'})$  are not tangential. Their intersection is transverse (which was the name used in [37], the preliminary release concerning these results).

Let us mention that if  $\text{Ker}(\mathcal{A})$  satisfies the deep-NSP with respect to the collection of models  $\mathcal{M}$  with constant  $\gamma$ , we have for all  $T' \in \text{Ker}(\mathcal{A})$  and all  $T \in P(\mathcal{M}^L) - P(\mathcal{M}^{L'})$  satisfying  $\|\mathcal{A}T\| \leq \varepsilon$

$$\|T'\| \leq \|T\| + \|T' - T\| \leq (\gamma + 1)\|T' - T\|.$$

Therefore,

$$\forall T' \in \text{Ker}(\mathcal{A}), \quad \|T'\| \leq (\gamma + 1)d_{loc}(T', P(\mathcal{M}^L) - P(\mathcal{M}^{L'})) \quad (31)$$

where we have set for any  $C \subset \mathbb{R}^{S^k}$

$$d_{loc}(T', C) = \inf_{T \in C, \|\mathcal{A}T\| \leq \varepsilon} \|T' - T\|.$$

The converse is also true, if  $\text{Ker}(\mathcal{A})$  satisfies (31), it satisfies the deep-NSP with respect to the collection of models  $\mathcal{M}$  with an appropriate constant. In the context of the usual compressed sensing (i.e., when  $K = 1$ ,  $\mathcal{M}^L$  contains  $L$ -sparse signals,  $\mathcal{A}$  is a rectangular matrix with full row rank and  $X$  is a vector), the localization appearing in  $d_{loc}$  can be discarded since the inequality must hold when  $T'$  is small and since in this case this localization has no effect. Therefore, in the compressed sensing context, (31) (and therefore deep-NSP) is the usual Null Space Property with respect to  $L$ -sparse vectors, as defined in [13]. However, deep-NSP is generalized to take into account deep factorization problems. This motivates the name.

In the general case, the deep-NSP can be understood as a local version of the generalized-NSP for  $\mathcal{A}$  relative to  $P(\cup_{L \in \mathbb{N}} \mathcal{M}^L) - P(\cup_{L \in \mathbb{N}} \mathcal{M}^L)$ , as defined in [7]. Our interest for the locality (as imposed by the constraint  $\|\mathcal{A}T\| \leq \varepsilon$ ) is motivated by the fact that we want to use the deep-NSP when the signal to noise ratio is controlled (i.e., the hypotheses of Theorem 1 are satisfied). Our stable recovery property therefore includes such hypotheses. Such locality hypotheses are needed to obtain Theorem 6.

Also, we have not adapted the robust-NSP defined in [7]. The benefit not to use this definition is to obtain a simpler definition for deep-NSP. In particular (30) does not involve the geometry of  $\mathcal{A}$  in the orthogonal complement of  $\text{Ker}(\mathcal{A})$ . Looking in detail at the benefit of this adaptation is of course, of a great interest.

Finally, notice that we trivially have the following two facts:

- For any  $\gamma' \geq \gamma$ : If  $\text{Ker}(\mathcal{A})$  satisfies the deep-NSP with respect to the collection of models  $\mathcal{M}$  with constant  $\gamma$ , then  $\text{Ker}(\mathcal{A})$  satisfies the deep-NSP with respect to the collection of models  $\mathcal{M}$  with constant  $\gamma'$ .
- For any  $\mathcal{M}' \subset \mathcal{M}$ : If  $\text{Ker}(\mathcal{A})$  satisfies the deep-NSP with respect to the collection of models  $\mathcal{M}$  with constant  $\gamma$ , then  $\text{Ker}(\mathcal{A})$  satisfies the deep-NSP with respect to the collection of models  $\mathcal{M}'$  with constant  $\gamma$ . In particular, if  $\text{Ker}(\mathcal{A})$  satisfies the deep-NSP with respect to the model  $\mathbb{R}^{S \times K}$  with constant  $\gamma$ , it satisfies the deep-NSP with respect to any collection of models, with constant  $\gamma$ .

**Theorem 5. Sufficient condition for stable recovery**

Assume  $\text{Ker}(\mathcal{A})$  satisfies the deep-NSP with respect to the collection of models  $\mathcal{M}$  and with the constant  $\gamma > 0$ . For any  $\mathbf{h}^*$  as in (28) with  $\eta$  and  $\delta$  (see (28) and (27)) sufficiently small, we have

$$\|P(\mathbf{h}^*) - P(\bar{\mathbf{h}})\| \leq \frac{\gamma}{\sigma_{min}} (\delta + \eta),$$

where  $\sigma_{\min}$  is the smallest non-zero singular value of  $\mathcal{A}$ . Moreover, if  $\bar{\mathbf{h}} \in \mathbb{R}_*^{S \times K}$

$$d_p([\mathbf{h}^*], [\bar{\mathbf{h}}]) \leq \frac{7(KS)^{\frac{1}{p}} \gamma}{\sigma_{\min}} \min \left( \|P(\bar{\mathbf{h}})\|_{\infty}^{\frac{1}{k}-1}, \|P(\mathbf{h}^*)\|_{\infty}^{\frac{1}{k}-1} \right) (\delta + \eta). \quad (32)$$

*Proof.* We have

$$\begin{aligned} \|\mathcal{A}(P(\mathbf{h}^*) - P(\bar{\mathbf{h}}))\| &\leq \|\mathcal{A}P(\mathbf{h}^*) - X\| + \|\mathcal{A}P(\bar{\mathbf{h}}) - X\| \\ &\leq \delta + \eta \end{aligned}$$

Geometrically, this means that  $P(\mathbf{h}^*)$  belongs to a cylinder centered at  $P(\bar{\mathbf{h}})$  whose direction is  $\text{Ker}(\mathcal{A})$  and whose section is defined using the operator  $\mathcal{A}$ . If we further decompose (the decomposition is unique)

$$P(\mathbf{h}^*) - P(\bar{\mathbf{h}}) = T + T',$$

where  $T' \in \text{Ker}(\mathcal{A})$  and  $T$  is orthogonal to  $\text{Ker}(\mathcal{A})$ , we have

$$\|\mathcal{A}(P(\mathbf{h}^*) - P(\bar{\mathbf{h}}))\| = \|\mathcal{A}T\| \geq \sigma_{\min} \|T\|, \quad (33)$$

where  $\sigma_{\min}$  is the smallest non-zero singular value of  $\mathcal{A}$ . We finally obtain

$$\|P(\mathbf{h}^*) - P(\bar{\mathbf{h}}) - T'\| = \|T\| \leq \frac{\delta + \eta}{\sigma_{\min}}.$$

The term on the left-hand side corresponds to the distance between a point in  $P(\mathcal{M}^{L^*}) - P(\mathcal{M}^{\bar{L}})$  (namely  $P(\mathbf{h}^*) - P(\bar{\mathbf{h}})$ ) and a point in  $\text{Ker}(\mathcal{A})$  (namely  $T'$ ).

Since  $\text{Ker}(\mathcal{A})$  satisfies the deep-NSP with constant  $\gamma$ , when  $\delta + \eta \leq \varepsilon$ , we obtain the first inequality of the theorem

$$\|P(\mathbf{h}^*) - P(\bar{\mathbf{h}})\| \leq \gamma \frac{\delta + \eta}{\sigma_{\min}}.$$

When  $\bar{\mathbf{h}} \in \mathbb{R}_*^{S \times K}$ , for  $\delta + \eta$  small, we can apply Theorem 1 and obtain (32).  $\square$

This proposition provides a sufficient condition to get stable recovery. The only significant hypothesis made on the factorization problem is that  $\text{Ker}(\mathcal{A})$  satisfies the deep-NSP with respect to the collection of models  $\mathcal{M}$ . One might ask whether this hypothesis is sharp or not. As expressed in the next theorem, the answer to this question is positive.

**Theorem 6. Necessary condition for stable recovery**

*Assume the stable recovery property holds: There exists  $C$  and  $\delta > 0$  such that for any  $\bar{L} \in \mathbb{N}$ ,  $\bar{\mathbf{h}} \in \mathcal{M}^{\bar{L}}$ , any  $X = \mathcal{A}P(\bar{\mathbf{h}}) + e$ , with  $\|e\| \leq \delta$ , any  $L^* \in \mathbb{N}$  and any  $\mathbf{h}^* \in \mathcal{M}^{L^*}$  such that*

$$\|\mathcal{A}P(\mathbf{h}^*) - X\|^2 \leq \|e\|$$

*we have*

$$d_2([\mathbf{h}^*], [\bar{\mathbf{h}}]) \leq C \min \left( \|P(\bar{\mathbf{h}})\|_{\infty}^{\frac{1}{k}-1}, \|P(\mathbf{h}^*)\|_{\infty}^{\frac{1}{k}-1} \right) \|e\|.$$

*Then,  $\text{Ker}(\mathcal{A})$  satisfies the deep-NSP with respect to the collection of models  $\mathcal{M}$  with constant*

$$\gamma = CS^{\frac{K-1}{2}} \sqrt{K} \sigma_{\max}$$

*(and  $\varepsilon = \delta$ ), where  $\sigma_{\max}$  is the spectral radius of  $\mathcal{A}$ .*



*Proof.* Let  $\bar{L}$  and  $\bar{L}' \in \mathbb{N}$  and  $\bar{\mathbf{h}} \in \mathcal{M}^{\bar{L}}$  and  $\bar{\mathbf{h}}' \in \mathcal{M}^{\bar{L}'}$  be such that  $\|\mathcal{A}(P(\bar{\mathbf{h}}) - P(\bar{\mathbf{h}}'))\| \leq \delta$ . We also consider throughout the proof  $T' \in \text{Ker}(\mathcal{A})$ . We assume that  $\|P(\bar{\mathbf{h}})\|_\infty \leq \|P(\bar{\mathbf{h}}')\|_\infty$ . If it is not the case, we simply switch  $\bar{\mathbf{h}}$  and  $\bar{\mathbf{h}}'$  in the definition of  $X$  and  $e$  below. We denote

$$X = \mathcal{A}P(\bar{\mathbf{h}}) \quad \text{and} \quad e = \mathcal{A}P(\bar{\mathbf{h}}) - \mathcal{A}P(\bar{\mathbf{h}}').$$

We have  $X = \mathcal{A}P(\bar{\mathbf{h}}') + e$  and  $\|e\| \leq \delta$ . Therefore, the hypothesis of the theorem (applied with  $\mathbf{h}^* = \bar{\mathbf{h}}$  and  $L^* = \bar{L}$ ) guarantees that

$$d_2([\bar{\mathbf{h}}], [\bar{\mathbf{h}}']) \leq C \|P(\bar{\mathbf{h}}')\|_\infty^{\frac{1}{k}-1} \|e\|.$$

Using the fact that  $e = \mathcal{A}P(\bar{\mathbf{h}}) - \mathcal{A}P(\bar{\mathbf{h}}')$  and  $T' \in \text{Ker}(\mathcal{A})$  we obtain

$$\|e\| = \|\mathcal{A}(P(\bar{\mathbf{h}}) - P(\bar{\mathbf{h}}') - T')\| \leq \sigma_{\max} \|P(\bar{\mathbf{h}}) - P(\bar{\mathbf{h}}') - T'\|.$$

where  $\sigma_{\max}$  is the spectral radius of  $\mathcal{A}$ . Therefore

$$d_2([\bar{\mathbf{h}}], [\bar{\mathbf{h}}']) \leq C \|P(\bar{\mathbf{h}}')\|_\infty^{\frac{1}{k}-1} \sigma_{\max} \|P(\bar{\mathbf{h}}) - P(\bar{\mathbf{h}}') - T'\|.$$

Finally, using Theorem 2 and the fact that  $\|P(\bar{\mathbf{h}})\|_\infty \leq \|P(\bar{\mathbf{h}}')\|_\infty$ , we obtain

$$\begin{aligned} \|P(\bar{\mathbf{h}}') - P(\bar{\mathbf{h}})\| &\leq S^{\frac{K-1}{2}} K^{1-\frac{1}{2}} \|P(\bar{\mathbf{h}}')\|_\infty^{1-\frac{1}{k}} d_2([\bar{\mathbf{h}}'], [\bar{\mathbf{h}}]) \\ &\leq CS^{\frac{K-1}{2}} \sqrt{K} \sigma_{\max} \|P(\bar{\mathbf{h}}) - P(\bar{\mathbf{h}}') - T'\| \\ &= \gamma \|P(\bar{\mathbf{h}}) - P(\bar{\mathbf{h}}') - T'\| \end{aligned}$$

for  $\gamma = CS^{\frac{K-1}{2}} \sqrt{K} \sigma_{\max}$ .

Summarizing, we conclude that under the hypothesis of the theorem: For any  $T \in P(\mathcal{M}^{\bar{L}}) - P(\mathcal{M}^{\bar{L}'})$  such that  $\|\mathcal{A}T\| \leq \delta$  we have for any  $T' \in \text{Ker}(\mathcal{A})$

$$\|T\| \leq \gamma \|T - T'\|.$$

□

## 7 Application to convolutional network

We consider a convolutional network as depicted in Figure 1. The network typically aims at performing a linear analysis or synthesis of a signal living in  $\mathbb{R}^N$ . The considered convolutional network is defined from a rooted directed acyclic graph  $\mathcal{G}(\mathcal{E}, \mathcal{N})$  composed of nodes  $\mathcal{N}$  and edges  $\mathcal{E}$ . Each edge connects two nodes. The root of the graph is denoted by  $r$  and the set containing all its leaves is denoted by  $\mathcal{F}$ . We denote by  $\mathcal{P}$  the set of all paths connecting the leaves and the root. We assume, without loss of generality, that the length of any path between any leaf and the root is independent of the considered leaf and equal to some constant  $K \geq 0$ . We also assume that, for any edge  $e \in \mathcal{E}$ , the number of edges separating  $e$  and the root is the same for all paths between  $e$  and  $r$ . It is called the depth of  $e$ . We also say that  $e$  belongs to the layer  $k$ . For any  $k \in \mathbb{N}_K$ , we denote the set containing all the edges of depth  $k$ , by  $\mathcal{E}(k)$ .

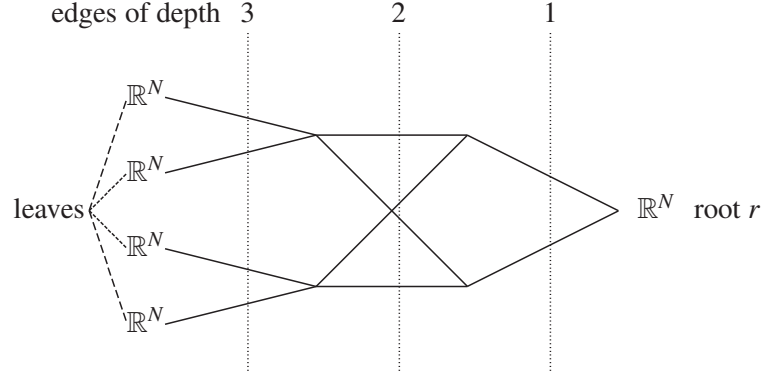


Figure 1: Example of the considered convolutional network. To every edge is attached a convolution kernel. The network does not involve non-linearities or sampling.

Moreover, to any edge  $e$  is attached a convolution kernel of support  $\mathcal{S}_e \subset \mathbb{N}_N$ . We assume (without loss of generality) that  $\sum_{e \in \mathcal{E}(k)} |\mathcal{S}_e|$  is independent of  $k$  ( $|\mathcal{S}_e|$  denotes the cardinality of  $\mathcal{S}_e$ ). We take

$$S = \sum_{e \in \mathcal{E}(1)} |\mathcal{S}_e|.$$

For any edge  $e$ , we consider the mapping  $\mathcal{T}_e : \mathbb{R}^S \rightarrow \mathbb{R}^N$  that maps any  $h \in \mathbb{R}^S$  into the convolution kernel  $h_e$ , attached to the edge  $e$ , whose support is  $\mathcal{S}_e$ . It simply writes at the right location (i.e. those in  $\mathcal{S}_e$ ) the entries of  $h$  defining the kernel on the edge  $e$ .

At each layer  $k$ , the convolutional network computes, for all  $e \in \mathcal{E}(k)$ , the convolution between the signal at the origin of  $e$ ; then, it attaches to any ending node the sum of all the convolutions arriving at that node. Examples of such convolutional networks includes wavelets, wavelet packets [38] or the fast transforms optimized in [10, 11]. It is clear that the operation performed at any layer depends linearly on the parameters  $h \in \mathbb{R}^S$  and that its results serves as inputs for the next layer. The convolutional network therefore depends on parameters  $\mathbf{h} \in \mathbb{R}^{S \times K}$  and takes the form

$$X = M_1(\mathbf{h}_1) \dots M_K(\mathbf{h}_K),$$

where the operators  $M_k$  satisfy the hypothesis of the present paper.

This section aims at identifying conditions such that any unknown parameters  $\bar{\mathbf{h}} \in \mathbb{R}^{S \times K}$  can be identified or stably recovered from  $X = M_1(\bar{\mathbf{h}}_1) \dots M_K(\bar{\mathbf{h}}_K)$  (possibly corrupted by an error).

In order to do so, let us define a few notations. Notice first that, we apply the convolutional network to an input  $x \in \mathbb{R}^{N|\mathcal{F}|}$ , where  $x$  is the concatenation of the signals  $x^f \in \mathbb{R}^N$  for  $f \in \mathcal{F}$ . Therefore,  $X$  is the (horizontal) concatenation of  $|\mathcal{F}|$  matrices  $X^f \in \mathbb{R}^{N \times N}$  such that

$$Xx = \sum_{f \in \mathcal{F}} X^f x^f \quad , \text{ for all } x \in \mathbb{R}^{N|\mathcal{F}|}.$$

Let us consider the convolutional network defined by  $\mathbf{h} \in \mathbb{R}^{S \times K}$  as well as  $f \in \mathcal{F}$  and  $n \in \mathbb{N}_N$ . The column of  $X$  corresponding to the entry  $n$  in the leaf  $f$  is the translation by  $n$  of

$$\sum_{p \in \mathcal{P}(f)} \mathcal{T}^p(\mathbf{h}) \tag{34}$$

where  $\mathcal{P}(f)$  contains all the paths of  $\mathcal{P}$  starting from the leaf  $f$  and

$$\mathcal{T}^p(\mathbf{h}) = \mathcal{T}_{e^1}(\mathbf{h}_1) * \dots * \mathcal{T}_{e^K}(\mathbf{h}_K) \quad , \text{ where } p = (e^1, \dots, e^K).$$

Moreover, we define for any  $k \in \mathbb{N}_K$  the mapping  $\mathbf{e}_k : \mathbb{N}_S \longrightarrow \mathcal{P}(k)$  which provides for any  $i \in \mathbb{N}_S$  the unique edge of  $\mathcal{P}(k)$  such that the  $i^{\text{th}}$  entry of  $h \in \mathbb{R}^S$  contributes to  $\mathcal{T}_{\mathbf{e}_k(i)}(h)$ . Also, for any  $\mathbf{i} \in \mathbb{N}_S^K$ , we denote  $\mathbf{p}_i = (\mathbf{e}_1(\mathbf{i}_1), \dots, \mathbf{e}_K(\mathbf{i}_K))$  and

$$\mathbf{I} = \{\mathbf{i} \in \mathbb{N}_S^K \mid \mathbf{p}_i \in \mathcal{P}\}.$$

The latter contains all the indices corresponding to a valid path in the network. For any set of parameters  $\mathbf{h} \in \mathbb{R}^{S \times K}$  and any path  $\mathbf{p} \in \mathcal{P}$ , we also denote by  $\mathbf{h}^{\mathbf{p}}$  the restriction of  $\mathbf{h}$  to its indices contributing to the kernels on the path  $\mathbf{p}$ . We also denote  $\mathbf{1} \in \mathbb{R}^S$  a vector of size  $S$  with all its entries equal to 1. For any edge  $e$ ,  $\mathbf{1}^e \in \mathbb{R}^S$  consists of zeroes except for the entries corresponding to the edge  $e$  which are equal to 1. For any  $\mathbf{p} = (e^1, \dots, e^K) \in \mathcal{P}$ , the support of  $M_1(\mathbf{1}^{e^1}) \dots M_K(\mathbf{1}^{e^K})$  is denoted by  $\mathcal{D}^{\mathbf{p}}$ .

Finally, we remind that because of Corollary 1, there exists a unique mapping

$$\mathcal{A} : \mathbb{R}^{S^K} \longrightarrow \mathbb{R}^{N \times N \mid \mathcal{F}}$$

such that

$$\mathcal{A}P(\mathbf{h}) = M_1(\mathbf{h}_1) \dots M_K(\mathbf{h}_K) \quad , \text{ for all } \mathbf{h} \in \mathbb{R}^{S \times K},$$

where  $P$  is the Segre embedding (defined in (5)).

**Proposition 8. Necessary condition of identifiability of a network**

- Either all the entries of  $M_1(\mathbf{1}) \dots M_K(\mathbf{1})$  belong to  $\{0, 1\}$  and then
  1. for any distinct  $\mathbf{p}$  and  $\mathbf{p}' \in \mathcal{P}$ , we have  $\mathcal{D}^{\mathbf{p}} \cap \mathcal{D}^{\mathbf{p}'} = \emptyset$ .
  2.  $\text{Ker}(\mathcal{A}) = \{T \in \mathbb{R}^{S^K} \mid \forall \mathbf{i} \in \mathbf{I}, T_{\mathbf{i}} = 0\}$ .
- or some of the entries of  $M_1(\mathbf{1}) \dots M_K(\mathbf{1})$  do not belong to  $\{0, 1\}$  and then  $\mathbb{R}^{S \times K}$  is not identifiable.

*Proof.* Throughout the proof, we define, for any  $\mathbf{i} \in \mathbb{N}_S^K$ ,  $\mathbf{h}^{\mathbf{i}} \in \mathbb{R}^{S \times K}$  by

$$\mathbf{h}_{k,j}^{\mathbf{i}} = \begin{cases} 1 & , \text{ if } j = \mathbf{i}_k \\ 0 & \text{ otherwise} \end{cases} \quad , \text{ for all } k \in \mathbb{N}_K \text{ and } j \in \mathbb{N}_S. \quad (35)$$

This notation shall not be confused with  $\mathbf{h}^{\mathbf{p}}$ , with  $\mathbf{p} \in \mathcal{P}$ .

- Let us first prove the first statement:  
We can easily check that  $(P(\mathbf{h}^{\mathbf{i}}))_{\mathbf{i} \in \mathbf{I}}$  forms a basis of  $\{T \in \mathbb{R}^{S^K} \mid \forall \mathbf{i} \in \mathbf{I}, T_{\mathbf{i}} = 0\}$ . We can also easily check using (34) that, for any  $\mathbf{i} \notin \mathbf{I}$ ,

$$\mathcal{A}P(\mathbf{h}^{\mathbf{i}}) = M_1(\mathbf{h}_1^{\mathbf{i}}) \dots M_K(\mathbf{h}_K^{\mathbf{i}}) = 0.$$

Therefore,  $\{T \in \mathbb{R}^{S^K} \mid \forall \mathbf{i} \in \mathbf{I}, T_{\mathbf{i}} = 0\} \subset \text{Ker}(\mathcal{A})$ .

Conversely, for any  $\mathbf{i} \in \mathbf{I}$ , we can deduce from (34) and the hypotheses of the proposition that all the entries of  $\mathcal{A}P(\mathbf{h}^{\mathbf{i}})$  are in  $\{0, 1\}$ . We denote  $D_{\mathbf{i}} = \{(i, j) \in \mathbb{N}_N \times \mathbb{N}_{N \mid \mathcal{F}} \mid \mathcal{A}P(\mathbf{h}^{\mathbf{i}})_{i,j} = 1\}$ . Using (again) the hypothesis of the proposition and (34), we

can prove that, for any distinct  $\mathbf{i}$  and  $\mathbf{j} \in \mathbf{I}$ , we have  $D_{\mathbf{i}} \cap D_{\mathbf{j}} = \emptyset$ . This easily leads to the item 1 of the first statement. We also deduce that

$$\text{rk}(\mathcal{A}) \geq |\mathbf{I}| = S^K - \dim(\{T \in \mathbb{R}^{S^K} \mid \forall \mathbf{i} \in \mathbf{I}, T_{\mathbf{i}} = 0\}).$$

Finally, we deduce that  $\dim(\text{Ker}(\mathcal{A})) \leq \dim(\{T \in \mathbb{R}^{S^K} \mid \forall \mathbf{i} \in \mathbf{I}, T_{\mathbf{i}} = 0\})$  and therefore

$$\text{Ker}(\mathcal{A}) = \{T \in \mathbb{R}^{S^K} \mid \forall \mathbf{i} \in \mathbf{I}, T_{\mathbf{i}} = 0\}.$$

- Let us now prove the second statement:

Using the hypothesis of the second statement and (34), we know that there is  $f \in \mathcal{F}$  and  $n \in \mathbb{N}_N$  such that

$$\sum_{p \in \mathcal{P}(f)} \mathcal{T}^p(1)_n \geq 2.$$

As a consequence, there is  $\mathbf{i}$  and  $\mathbf{j} \in \mathbb{N}_S^K$  with  $\mathbf{i} \neq \mathbf{j}$  and

$$\mathcal{T}^{p_{\mathbf{i}}}(\mathbf{h}^{\mathbf{i}})_n = \mathcal{T}^{p_{\mathbf{j}}}(\mathbf{h}^{\mathbf{j}})_n = 1.$$

Therefore,

$$\mathcal{A}P(\mathbf{h}^{\mathbf{i}}) = \mathcal{A}P(\mathbf{h}^{\mathbf{j}})$$

and the network is not identifiable. □

The interest of the condition in Proposition 8 is that it can easily be computed when  $N \times N|\mathcal{F}|$  is not too large. Notice that, beside the known examples in blind-deconvolution (i.e. when  $K = 2$  and  $|\mathcal{P}| = 1$ ) [2, 6], there are known (truly deep) convolutional networks that satisfy the condition of the first statement of Proposition 8. For instance, the convolutional network corresponding to the un-decimated Haar (wavelet)<sup>2</sup> transform is a tree and for any of its leaves  $f \in \mathcal{F}$ ,  $|\mathcal{P}(f)| = 1$ . Moreover, the support of the kernel living on the edge  $e$ , of depth  $k$ , on this path is  $\{0, 2^k\}$ . The first condition of Proposition 8 therefore holds.

We also have the following proposition.

**Proposition 9.** *If  $|\mathcal{P}| = 1$  and all the entries of  $M_1(1) \dots M_K(1)$  belong to  $\{0, 1\}$ , then  $\text{Ker}(\mathcal{A}) = \{0\}$  and  $\text{Ker}(\mathcal{A})$  satisfies the deep-NSP with respect to any model collection  $\mathcal{M}$  with constant  $\gamma = 1$ . Moreover, we have  $\sigma_{\min} = \sqrt{N}$ .*

*Proof.* The fact that, under the hypotheses of the proposition,  $\text{Ker}(\mathcal{A}) = \{0\}$  is a direct consequence of Proposition 8. The deep-NSP property and the value of  $\gamma$  also immediately follow from the definition of the deep-NSP.

To calculate  $\sigma_{\min}$ , let us consider  $T \in \mathbb{R}^{S^K}$  and express it under the form  $T = \sum_{\mathbf{i} \in \mathbf{I}} T_{\mathbf{i}} P(\mathbf{h}^{\mathbf{i}})$ , where  $\mathbf{h}^{\mathbf{i}}$  is defined (35). Let us also remind that, applying Proposition 8, the supports of  $\mathcal{A}P(\mathbf{h}^{\mathbf{i}})$  and  $\mathcal{A}P(\mathbf{h}^{\mathbf{j}})$  are disjoint, when  $\mathbf{i} \neq \mathbf{j}$ . Let us finally add that, since  $\mathcal{A}P(\mathbf{h}^{\mathbf{j}})$  is the matrix of a convolution with a Dirac mass, its support is of size  $N$ . We finally have

$$\begin{aligned} \|\mathcal{A}T\|^2 &= \left\| \sum_{\mathbf{i} \in \mathbf{I}} T_{\mathbf{i}} \mathcal{A}P(\mathbf{h}^{\mathbf{i}}) \right\|^2, \\ &= N \sum_{\mathbf{i} \in \mathbf{I}} T_{\mathbf{i}}^2 = N \|T\|^2, \end{aligned}$$

from which we deduce the value of  $\sigma_{\min}$ . □

---

<sup>2</sup>Un-decimated means computed with the "Algorithme à trous", [38], Section 5.5.2 and 6.3.2. The Haar wavelet is described in [38], Section 7.2.2, p. 247 and Example 7.7, p. 235

In the sequel, we establish stability results for a convolutional network estimator. In order to do so, we consider a convolutional network of known structure  $\mathcal{G}(\mathcal{E}, \mathcal{N})$  and  $(S_e)_{e \in \mathcal{E}}$  but defined by unknown parameters  $\bar{\mathbf{h}} \in \mathbb{R}^{S \times K}$ . We consider the noisy situation where

$$X = M_1(\bar{\mathbf{h}}_1) \dots M_K(\bar{\mathbf{h}}_K) + e,$$

with  $\|e\| \leq \delta$  and an estimate  $\mathbf{h}^* \in \mathbb{R}^{S \times K}$  such that

$$\|M_1(\mathbf{h}_1^*) \dots M_K(\mathbf{h}_K^*) - X\| \leq \eta.$$

We say that two networks sharing the same structure and defined by  $\mathbf{h}$  and  $\mathbf{g} \in \mathbb{R}^{S \times K}$  are equivalent if and only if

$$\forall \mathbf{p} \in \mathcal{P}, \exists (\lambda_e)_{e \in \mathbf{p}} \in \mathbb{R}^{\mathbf{p}}, \text{ such that } \prod_{e \in \mathbf{p}} \lambda_e = 1 \text{ and } \forall e \in \mathbf{p}, \mathcal{T}_e(\mathbf{g}) = \lambda_e \mathcal{T}_e(\mathbf{h}).$$

The equivalence class of  $\mathbf{h} \in \mathbb{R}^{S \times K}$  is denoted by  $\{\mathbf{h}\}$ . For any  $p \in [1, +\infty]$ , we define

$$\delta_p(\{\mathbf{h}\}, \{\mathbf{g}\}) = \left( \sum_{\mathbf{p} \in \mathcal{P}} d_p([\mathbf{h}^{\mathbf{p}}], [\mathbf{g}^{\mathbf{p}}])^p \right)^{\frac{1}{p}},$$

where we remind that  $\mathbf{h}^{\mathbf{p}}$  (resp  $\mathbf{g}^{\mathbf{p}}$ ) denotes the restriction of  $\mathbf{h}$  (resp  $\mathbf{g}$ ) to the path  $\mathbf{p}$  and  $d_p$  is defined in Definition 1. Since  $d_p$  is a metric, we easily prove that  $\delta_p$  is a metric between network classes.

**Theorem 7.** *If all the entries of  $M_1(1) \dots M_K(1)$  belong to  $\{0, 1\}$  and if there exists  $\varepsilon > 0$  such that for all  $e \in \mathcal{E}$ ,  $\|\mathcal{T}_e(\bar{\mathbf{h}})\|_\infty \geq \varepsilon$  then*

$$\delta_p(\{\mathbf{h}^*\}, \{\bar{\mathbf{h}}\}) \leq 7(KS)^{\frac{1}{p}} \varepsilon^{1-K} \frac{\delta + \eta}{\sqrt{N}}.$$

*Proof.* Let us consider a path  $\mathbf{p} \in \mathcal{P}$ , using (34), since all the entries of  $M_1(1) \dots M_K(1)$  belong to  $\{0, 1\}$ , all the entries of  $M_1(1^{\mathbf{p}}) \dots M_K(1^{\mathbf{p}})$  belong to  $\{0, 1\}$ . Therefore, we can apply Proposition 9 and Theorem 5 to the restriction of the convolutional network to  $\mathbf{p}$  and obtain

$$d_p([\mathbf{h}^{\mathbf{p}}], [\bar{\mathbf{h}}^{\mathbf{p}}]) \leq \frac{7(KS)^{\frac{1}{p}}}{\sqrt{N}} \min \left( \|P(\bar{\mathbf{h}}^{\mathbf{p}})\|_\infty^{\frac{1}{K}-1}, \|P([\mathbf{h}^*]^{\mathbf{p}})\|_\infty^{\frac{1}{K}-1} \right) (\delta^{\mathbf{p}} + \eta^{\mathbf{p}}),$$

where  $\delta^{\mathbf{p}}$  and  $\eta^{\mathbf{p}}$  are the restrictions of the errors on  $\mathcal{D}^{\mathbf{p}}$ .

We therefore have

$$d_p([\mathbf{h}^{\mathbf{p}}], [\bar{\mathbf{h}}^{\mathbf{p}}]) \leq \frac{7(KS)^{\frac{1}{p}}}{\sqrt{N}} \varepsilon^{1-K} (\delta^{\mathbf{p}} + \eta^{\mathbf{p}}),$$

and finally

$$\begin{aligned} \delta_p(\{\mathbf{h}^*\}, \{\bar{\mathbf{h}}\}) &\leq \frac{7(KS)^{\frac{1}{p}} \varepsilon^{1-K}}{\sqrt{N}} \left( \sum_{\mathbf{p} \in \mathcal{P}} (\delta^{\mathbf{p}} + \eta^{\mathbf{p}})^p \right)^{\frac{1}{p}}, \\ &\leq \frac{7(KS)^{\frac{1}{p}} \varepsilon^{1-K}}{\sqrt{N}} (\delta + \eta). \end{aligned}$$

□

## Acknowledgment

Joseph Landsberg is supported by NSF DMS-1405348.

## References

- [1] Alekh Agarwal, Animashree Anandkumar, and Praneeth Netrapalli. Exact recovery of sparsely used overcomplete dictionaries. *stat*, 1050:8, 2013.
- [2] Arif Ahmed, Benjamin Recht, and Justin Romberg. Blind deconvolution using convex programming. *IEEE Transactions on Information Theory*, 60(3):1711–1732, 2014.
- [3] Sanjeev Arora, Aditya Bhaskara, Rong Ge, and Tengyu Ma. Provable bounds for learning some deep representations. In *ICML*, pages 584–592, 2014.
- [4] Sanjeev Arora, Rong Ge, Ravindran Kannan, and Ankur Moitra. Computing a non-negative matrix factorization—provably. In *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*, pages 145–162. ACM, 2012.
- [5] Sanjeev Arora, Rong Ge, and Ankur Moitra. New algorithms for learning incoherent and overcomplete dictionaries. In *COLT*, pages 779–806, 2014.
- [6] Sohail Bahmani and Justin Romberg. Lifting for blind deconvolution in random mask imaging: Identifiability and convex relaxation. *SIAM Journal on Imaging Sciences*, 8(4):2203–2238, 2015.
- [7] Anthony Bourrier, Mike Davies, Tomer Peleg, Patrick Pérez, and Rémi Gribonval. Fundamental performance limits for ideal decoders in high-dimensional linear inverse problems. *IEEE Transactions on Information Theory*, 60(12):7928–7946, 2014.
- [8] Emmanuel Candès, Justin Romberg, and Terence Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on information theory*, 52(2):489–509, 2006.
- [9] Emmanuel J Candès, Thomas Strohmer, and Vladislav Voroninski. Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming. *Communications on Pure and Applied Mathematics*, 66(8):1241–1274, 2013.
- [10] Olivier Chabiron, François Malgouyres, Jean-Yves Tourneret, and Nicolas Dobigeon. Toward fast transform learning. *International Journal of Computer Vision*, pages 1–22, 2014.
- [11] Olivier Chabiron, François Malgouyres, Herwig Wendt, and Jean-Yves Tourneret. Optimization of a fast transform structured as a convolutional tree. *preprint HAL*, (hal-01258514), 2016.
- [12] Sunav Choudhary and Urbashi Mitra. Identifiability scaling laws in bilinear inverse problems. *arXiv preprint arXiv:1402.2637*, 2014.
- [13] Albert Cohen, Wolfgang Dahmen, and Ronald DeVore. Compressed sensing and best-term approximation. *Journal of the American mathematical society*, 22(1):211–231, 2009.
- [14] Ron A. DeVore. Nonlinear approximation. *Acta Numerica*, 7:51–150, 1998.
- [15] David L. Donoho. Compressed sensing. *IEEE Transactions on information theory*, 52(4):1289–1306, 2006.

- [16] David L Donoho and Victoria Stodden. When does non-negative matrix factorization give a correct decomposition into parts? 2003.
- [17] Quan Geng, Huan Wangy, and John Wright. On the local correctness of  $\ell_1$ -minimization for dictionary learning. In *International Symposium on Information Theory (ISIT)*, 2014.
- [18] Wallace Givens. Computation of plain unitary rotations transforming a general matrix to triangular form. *Journal of the Society for Industrial and Applied Mathematics*, 6(1):26–50, 1958.
- [19] Rémi Gribonval, Rodolphe Jenatton, and Francis Bach. Sample complexity of dictionary learning and other matrix factorizations. *Information Theory, IEEE Transactions on*, 61(6):3469–3486, June 2015.
- [20] Rémi Gribonval and Karin Schnass. Dictionary identification - sparse matrix-factorisation via  $\ell_1$ -minimisation. *IEEE transaction on information theory*, 56(7):3523–3539, 2010.
- [21] Benjamin D Haeffele and René Vidal. Global optimality in tensor factorization, deep learning, and beyond. *arXiv preprint arXiv:1506.07540*, 2015.
- [22] Joe Harris. *Algebraic geometry*, volume 133 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995. A first course, Corrected reprint of the 1992 original.
- [23] Jonathan D. Hauenstein and Andrew J. Sommese. Witness sets of projections. *Applied Mathematics and Computation*, 217(7):3349 – 3354, 2010.
- [24] Jonathan D. Hauenstein and Andrew J. Sommese. Membership tests for images of algebraic sets by linear projections. *Applied Mathematics and Computation*, 219(12):6809 – 6818, 2013.
- [25] Rodolphe Jenatton, Rémi Gribonval, and Francis Bach. Local stability and robustness of sparse dictionary learning in the presence of noise. *arxiv*, 2012.
- [26] Risi Kondor, Nedelina Teneva, and Vikas Garg. Multiresolution matrix factorization. In *Proceedings of the 31st International Conference on Machine Learning (ICML-14)*, pages 1620–1628, 2014.
- [27] Joseph M. Landsberg. *Tensors: Geometry and Applications*, volume 128 of *Graduate studies in mathematics*. American Mathematical Soc., 2012.
- [28] Hans Laurberg, Mads Græsbøll Christensen, Mark D Plumbley, Lars Kai Hansen, and Søren Holdt Jensen. Theorems on positive data: On the uniqueness of nmf. *Computational intelligence and neuroscience*, 2008, 2008.
- [29] Daniel D Lee and H Sebastian Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755):788–791, 1999.
- [30] Xiaodong Li, Shuyang Ling, Thomas Strohmer, and Ke Wei. Rapid, robust, and reliable blind deconvolution via nonconvex optimization. *CoRR*, abs/1606.04933, 2016.
- [31] Shuyang Ling and Thomas Strohmer. Blind deconvolution meets blind demixing: Algorithms and performance bounds. *arXiv preprint arXiv:1512.07730*, 2015.
- [32] Shuyang Ling and Thomas Strohmer. Self-calibration and biconvex compressive sensing. *Inverse Problems*, 31(11):115002, 2015.
- [33] Siwei Lyu and Xin Wang. On algorithms for sparse multi-factor nmf. In *Proceedings of the 26th International Conference on Neural Information Processing Systems, NIPS’13*, pages 602–610, USA, 2013. Curran Associates Inc.

- [34] Luc Le Magoarou. *Matrices efficaces pour le traitement du signal et l'apprentissage automatique*. PhD thesis, Université Bretagne Loire, 2016.
- [35] Luc Le Magoarou and Rémi Gribonval. Are there approximate fast fourier transforms on graphs? In *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 4811–4815, 2016.
- [36] Luc Le Magoarou and Rémi Gribonval. Flexible multi-layer sparse approximations of matrices and applications. *IEEE Journal of Selected Topics in Signal Processing*, 10(4):688–700, 2016.
- [37] François Malgouyres and Joseph Landsberg. On the identifiability and stable recovery of deep/multi-layer structured matrix factorization. In *IEEE, Info. Theory Workshop*, Sept. 2016.
- [38] Stéphane Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, Boston, 1998.
- [39] Behnam Neyshabur and Rina Panigrahy. Sparse matrix factorization. *arXiv preprint arXiv:1311.3315*, 2013.
- [40] Ron Rubinstein, Alfred Bruckstein, and Michael Elad. Dictionaries for sparse representation modeling. *Proc. IEEE - Special issue on applications of sparse representation and compressive sensing*, 98(6):1045–1057, 2010.
- [41] Cristian Rusu, Nuria González-Prelcic, and Robert W Heath. Fast orthonormal sparsifying transforms based on householder reflectors. *IEEE Transactions on Signal Processing*, 64(24):6589–6599, 2016.
- [42] Cristian Rusu and John Thompson. Learning fast sparsifying transforms. *arXiv preprint arXiv:1611.08230*, 2016.
- [43] Karin Schnass. On the identifiability of overcomplete dictionaries via the minimisation principle underlying k-svd. *Applied and Computational Harmonic Analysis*, 37(3):464–491, 2014.
- [44] Igor R. Shafarevich. *Basic algebraic geometry. 1*. Springer, Heidelberg, third edition, 2013. Varieties in projective space.
- [45] Daniel Spielmana, Huan Wang, and John Wright. Exact recovery of sparsely-used dictionaries. In *COLT*, pages 37–1, 2012.