Blind Multilinear Identification

Lek-Heng Lim and Pierre Comon, *Fellow, IEEE*

***Abstract*— We discuss a technique that allows blind recovery of signals or blind identification of mixtures in instances where such recovery or identification were previously thought to be impossible. These instances include: 1) closely located or highly correlated sources in antenna array processing; 2) highly correlated spreading codes in code division multiple access (CDMA) radio communication; and 3) nearly dependent spectra in fluorescence spectroscopy. These have important implications. In the case of antenna array processing, it allows for joint localization and extraction of multiple sources from the measurement of a noisy mixture recorded on multiple sensors in an entirely deterministic manner. In the case of CDMA, it allows the possibility of having a number of users larger than the spreading gain. In the case of fluorescence spectroscopy, it allows for detection of nearly identical chemical constituents. The proposed technique involves the solution of a bounded coherence low-rank multilinear approximation problem. We show that bounded coherence allows us to establish existence and uniqueness of the recovered solution. We will provide some statistical motivation for the approximation problem and discuss greedy approximation bounds. To provide the theoretical underpinnings for this technique, we develop a corresponding theory of sparse separable decompositions of functions, including notions of rank and nuclear norm that can be specialized to the usual ones for matrices and operators and also be applied to hypermatrices and tensors.**

***Index Terms*— Source separation, array signal processing, system identification, channel estimation, remote sensing, fluorescence, function approximation, harmonic analysis, greedy algorithms, inverse problems.**

I. Introduction

THERE are two simple ideas for reducing the complexity or dimension of a problem that are widely applicable because of their simplicity and generality:

- **Sparsity:** resolving a complicated entity, represented by a function f, into a sum of a small number of simple or elemental constituents:

$$f = \sum_{p=1}^{r} \alpha_p g_p.$$

- **Separability:** decoupling a complicated entity, represented by a function g, that depends on multiple factors into a product of simpler constituents, each depending only on one factor:

$$g(\mathbf{x}_1, \dots, \mathbf{x}_d) = \prod_{k=1}^{d} \varphi_k(\mathbf{x}_k).$$

The two ideas underlie some of the most useful techniques in engineering and science — Fourier, wavelets, and other orthogonal or sparse representations of signals and images, singular value and eigenvalue decompositions of matrices, separation-of-variables, Fast Fourier Transform, mean field approximation, etc. This article examines the model that combines these two simple ideas:

$$f(\mathbf{x}_1, \dots, \mathbf{x}_d) = \sum_{p=1}^{r} \alpha_p \prod_{k=1}^{d} \varphi_{kp}(\mathbf{x}_k), \tag{1}$$

and we are primarily interested in its *inverse problem*, i.e., identification of the factors φ_{kp} based on noisy measurements of f. We shall see that this is a surprisingly effective method for a wide range of identification problems.

Here, f is approximately encoded by r scalars, $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_r) \in \mathbb{C}^r$, and dr functions, φ_{kp}, $k = 1, \dots, d$; $p = 1, \dots, r$. Since d and r are both assumed to be small, we expect (1) to be a very compact, possibly approximate, representation of f. We will assume that all these functions live in a Hilbert space with inner product $\langle \cdot, \cdot \rangle$, and that φ_{kp} are of unit norm (clearly possible since the norm of φ_{kp} can be 'absorbed into' the coefficient α_p in (1)).

Let $\mu_k = \max_{p \neq q} |\langle \varphi_{kp}, \varphi_{kq} \rangle|$ and define the *relative incoherence* $\omega_k = (1 - \mu_k)/\mu_k$ for $k = 1, \dots, d$. Note that $\mu_k \in [0, 1]$ and $\omega_k \in [0, \infty]$. We will show that if $d \geq 3$, and

$$\sum_{k=1}^{d} \omega_k \geq 2r - 1, \tag{2}$$

then the decomposition in (1) is essentially *unique* and *sparsest* possible, i.e., r is minimal. Hence we may in principle identify φ_{kp} based only on measurements of the mixture f.

One of the keys in the identifiability requirement is that $d \geq 3$ or otherwise (when $d = 1$ or 2) the result would not hold. We will show that the condition $d \geq 3$ however leads to a difficulty (that does not happen when $d = 1$ or 2). Since it is rarely, if not never, the case that one has the exact values of f, the decomposition (1) is only useful in an idealized scenario. In reality, one has $\hat{f} = f + \varepsilon$, an estimate of f corrupted by noise ε. Solving the inverse problem to (1) would require that we solve a best approximation problem. For example,

Manuscript received December 19, 2012; revised June 17, 2013; accepted September 4, 2013. Date of publication November 20, 2013; date of current version January 15, 2014. L.-H. Lim was supported in part by AFOSR Young Investigator Award FA9550-13-1-0133, in part by the National Science Foundation (NSF) Collaborative Research under Grant DMS 1209136, and in part by NSF CAREER Award DMS 1057064. P. Comon was supported by the European Research Council under the European Community's Seventh Framework Programme FP7/2007-2013 under Grant Agreement 320594.

L.-H. Lim is with the Computational and Applied Mathematics Initiative, Department of Statistics, University of Chicago, Chicago, IL 60637 USA (e-mail: lekheng@galton.uchicago.edu).

P. Comon is with the GIPSA-Laboratory, St. Martin d'Heres Cedex F-38402, France (e-mail: p.comon@ieee.org).

Communicated by G. Matz, Associate Editor for Detection and Estimation.
Digital Object Identifier 10.1109/TIT.2013.2291876

with the appropriate noise models (see Section V), the best approximation problem often takes the form

$$\operatorname*{argmin}_{\boldsymbol{\alpha}\in\mathbb{C}^r,\ \|\varphi_{kp}\|=1} \left\| \hat{f} - \sum_{p=1}^{r} \alpha_p \prod_{k=1}^{d} \varphi_{kp} \right\|, \tag{3}$$

with $\|\cdot\|$ an L^2-norm. Now, the trouble is that when $d \geq 3$, this best approximation problem may not have a solution — because the infimum of the loss function is unattainable in general, as we will discuss in Section VIII-A. In view of this, our next result is that when

$$\prod_{k=1}^{d}(1+\omega_k) > r, \tag{4}$$

the infimum in (3) is always attainable, thereby alleviating the aforementioned difficulty. A condition that meets both (2) and (4) follows from the arithmetic-geometric mean inequality

$$\left[\prod_{k=1}^{d}(1+\omega_k)\right]^{1/d} \leq 1 + \frac{1}{d}\sum_{k=1}^{d}\omega_k.$$

II. Sparse Separable Decompositions

The notion of *sparsity* dates back to harmonic analysis [52], [66], [74] and approximation theory [68], and has received a lot of recent attention in compressive sensing [15], [26], [29]. The notion of *separability* is also classical; it is the basis behind the separation-of-variables technique in partial differential equations [6] and special functions [55], fast Fourier transforms on arbitrary groups [53], mean field approximations in statistical physics [45], and the naïve Bayes model in machine learning [5], [48]. We describe a simple model that incorporates the two notions.

The function $f : X \to \mathbb{C}$ or $\mathbb{R}$ to be resolved into simpler entities will be referred to as our *target function*. We will treat the discrete (X is finite or countably infinite) and continuous (X is a continuum) cases on an equal footing. The discrete cases are when f is a vector (if $X = [n_1] := \{1, \dots, n_1\}$), a matrix (if $X = [n_1] \times [n_2]$), a hypermatrix (if $X = [n_1] \times [n_2] \times \cdots \times [n_d]$), while the usual continuous cases are when f is a function on some domain $X = \Omega \subseteq \mathbb{R}^m$ or $\mathbb{C}^m$. In the discrete cases, the set of target functions under consideration are identified with $\mathbb{C}^{n_1}$, $\mathbb{C}^{n_1\times n_2}$, $\mathbb{C}^{n_1\times n_2\times\cdots\times n_d}$ respectively whereas in the continuous cases, we usually impose some additional regularity structures such integrability or differentability, so that the set of target functions under consideration are $L^2(\Omega)$ or $C^\infty(\Omega)$ or $H^k(\Omega) = W^{k,2}(\Omega)$, etc. We will only assume that the space of target functions is a Hilbert space. Note that the requirement $d \geq 3$ implies that f is at least a 3-dimensional hypermatrix in the discrete case or a function of at least three continuous variables, i.e., $m \geq 3$, in the continuous case. The identifiability does not work for (usual 2-dimensional) matrices or bivariate functions. With (1) in mind, we will call f a *d-partite* or *multipartite* function if we wish to partition its arguments into d blocks of variables.

We will briefly examine the decompositions and approximations of our target function into a sum or integral of separable functions, adopting a tripartite notation for simplicity. There are three cases:

- **Continuous:**
$$f(\mathbf{x},\mathbf{y},\mathbf{z}) = \int_T \theta(\mathbf{x},\mathbf{t})\varphi(\mathbf{y},\mathbf{t})\psi(\mathbf{z},\mathbf{t})\,d\nu(\mathbf{t}). \tag{5}$$
Here, we assume that ν is some given Borel measure and that T is compact.
- **Semidiscrete:**
$$f(\mathbf{x},\mathbf{y},\mathbf{z}) = \sum_{p=1}^{r} \theta_p(\mathbf{x})\varphi_p(\mathbf{y})\psi_p(\mathbf{z}). \tag{6}$$
This may be viewed as a discretization of the continuous case in the $\mathbf{t}$ variable, i.e., $\theta_p(\mathbf{x}) = \theta(\mathbf{x},\mathbf{t}_p)$, $\varphi_p(\mathbf{y}) = \varphi(\mathbf{y},\mathbf{t}_p)$, $\psi_p(\mathbf{z}) = \psi(\mathbf{z},\mathbf{t}_p)$.
- **Discrete:**
$$a_{ijk} = \sum_{p=1}^{r} u_{ip}v_{jp}w_{kp}. \tag{7}$$
This may be viewed as a further discretization of the semidiscrete case, i.e. $a_{ijk} = f(\mathbf{x}_i,\mathbf{y}_j,\mathbf{z}_k)$, $u_{ip} = \theta_p(\mathbf{x}_i)$, $v_{jp} = \varphi_p(\mathbf{y}_j)$, $w_{kp} = \psi_p(\mathbf{z}_k)$.

It is clear that when i, j, k take finitely many values, the discrete decomposition (7) is always possible with a finite r since the space is of finite dimension. If i, j, k could take infinitely many values, then the finiteness of r requires that equality be replaced by approximation to any arbitrary precision $\varepsilon > 0$ in some suitable norm. This follows from the following observation about the semidiscrete decomposition: The space of functions with a semidiscrete representation as in (6), with r finite, is dense in $C^0(\Omega)$, the space of continuous functions. This is just a consequence of the Stone-Weierstrass theorem [20]. Discussion of the most general case (5) would require us to go into integral operators, which we will not do as in the present framework we are interested in applications that rely on the inverse problems corresponding to (6) and (7). Nonetheless (5) is expected to be useful and we state it here for completeness. Henceforth, we will simply refer to (6) or (7) as a *multilinear decomposition*, by which we mean a decomposition into a *linear combination of separable functions*. We note here that the finite-dimensional discrete version has been studied under several different names — see Section IX. Our emphasis in this paper is the semidiscrete version (6) that applies to multipartite functions on arbitrary domains and are not necessarily finite-dimensional. As such, we will frame most of our discussions in terms of the semidiscrete case, which of course includes the discrete version (7) as a special case (when $\mathbf{x}, \mathbf{y}, \mathbf{z}$ take only finite discrete values).

Example 1 (Mixture of Gaussians). Multilinear decompositions arise in many contexts. In machine learning or nonparametric statistics, a fact of note is that Gaussians are separable

$$\exp(x^2+y^2+z^2) = \exp(x^2)\exp(y^2)\exp(z^2).$$

More generally for symmetric positive-definite $A \in \mathbb{R}^{n\times n}$ with eigenvalues $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$,

$$\exp(\mathbf{x}^{\mathsf{T}}A\mathbf{x}) = \exp(\mathbf{z}^{\mathsf{T}}\Lambda\mathbf{z}) = \prod_{i=1}^{n}\exp(\lambda_i z_i^2),$$

under a linear change of coordinates $\mathbf{z} = Q^{\mathsf{T}}\mathbf{x}$ where $A = Q\Lambda Q^{\mathsf{T}}$. Hence Gaussian mixture models of the form

$$f(\mathbf{x}) = \sum_{j=1}^{m} \alpha_j \exp[(\mathbf{x} - \boldsymbol{\mu}_j)^{\mathsf{T}} A_j (\mathbf{x} - \boldsymbol{\mu}_j)],$$

where $A_i A_j = A_j A_i$ for all $i \neq j$ (and therefore $A_1, \ldots, A_m$ have a common eigenbasis) may likewise be transformed with a suitable linear change of coordinates into a multilinear decomposition as in (6).

We will later see several more examples from signal processing, telecommunications, and spectroscopy.

A. Modeling

The multilinear decomposition — an *additive* decomposition into *multiplicatively* decomposable components — is extremely simple but models a wide range of phenomena in signal processing and spectroscopy. The main message of this article is that the corresponding inverse problem — recovering the factors $\theta_p, \varphi_p, \psi_p$ from noisy measurements of f — can be solved under mild assumptions and yields a class of techniques for a range of applications (cf. Section IX) that we shall collectively call *multilinear identification*. We wish to highlight in particular that multilinear identification gives a deterministic approach for solving the problem of joint localization and estimation of radiating sources with short data lengths. This is superior to previous cumulants-based approaches [18], which require (i) longer data lengths; and (ii) statistically independent sources.

The experienced reader would probably guess that such a powerful technique must be fraught with difficulties and he would be right. The inverse problem to (6), like most other inverse problems, faces issues of existence, uniqueness, and computability. The approximation problem involved can be ill-posed in the worst possible way (cf. Section III). Fortunately, in part prompted by recent work in compressed sensing [9], [15], [26], [27], [33] and matrix completion [7], [8], [30], [34]), we show that mild assumptions on *coherence* allows one to overcome most of these difficulties (cf. Section VIII).

III. Finite Rank Multipartite Functions

In this section, we will discuss the notion of rank, which measures the sparsity of a multilinear decomposition, and the notion of Kruskal rank, which measures the uniqueness of a multilinear decomposition in a somewhat more restrictive sense. Why is uniqueness important? It can be answered in one word: Identifiability. More specifically, a unique decomposition means that we may in principle identify the factors. To be completely precise, we will first need to define the terms in the previous sentence, namely, 'unique', 'decomposition', and 'factor'. Before we do that, we will introduce the tensor product notation. It is not necessary to know anything about tensor product of Hilbert spaces to follow what we present below. We shall assume that all our Hilbert spaces are separable and so there is no loss of generality in assuming at the outset that they are just $L^2(X)$ for some σ-finite X.

Let $X_1, \ldots, X_d$ be σ-finite measurable spaces. There is a natural Hilbert space isomorphism

$$L^2(X_1 \times \cdots \times X_d) \cong L^2(X_1) \otimes \cdots \otimes L^2(X_d). \tag{8}$$

In other words, every d-partite L^2-function $f : X_1 \times \cdots \times X_d \to \mathbb{C}$ may be expressed as[1]

$$f(\mathbf{x}_1, \ldots, \mathbf{x}_d) = \sum_{p=1}^{\infty} \varphi_{1p}(\mathbf{x}_1) \cdots \varphi_{dp}(\mathbf{x}_d), \tag{9}$$

with $\varphi_{kp} \in L^2(X_k)$. The *tensor product* of functions $\varphi_1 \in L^2(X_1), \ldots, \varphi_d \in L^2(X_d)$ is denoted by $\varphi_1 \otimes \cdots \otimes \varphi_d$ and is the function in $L^2(X_1 \times \cdots \times X_d)$ defined by

$$\varphi_1 \otimes \cdots \otimes \varphi_d(\mathbf{x}_1, \ldots, \mathbf{x}_d) = \varphi_1(\mathbf{x}_1) \cdots \varphi_d(\mathbf{x}_d).$$

With this notation, we may rewrite (9) as

$$f = \sum_{p=1}^{\infty} \varphi_{1p} \otimes \cdots \otimes \varphi_{dp}.$$

A point worth noting here is that:

> "*Multipartite functions are infinite-dimensional tensors.*"

Finite-dimensional tensors are simply the special case when $X_1, \ldots, X_d$ are all finite sets (see Example 6). In particular, a multivariate function[2] $f \in L^2(\mathbb{R}^d)$ is a an infinite-dimensional tensor that can expressed as an infinite sum of a tensor product of $\varphi_{1p}, \ldots, \varphi_{dp} \in L^2(\mathbb{R})$ and $L^2(\mathbb{R}^d) \cong L^2(\mathbb{R}) \otimes \cdots \otimes L^2(\mathbb{R})$. We shall have more to say about this later in conjunction with Kolmogorov's superposition principle for multivariate functions.

In this paper, functions having a *finite* decomposition will play a central role; for these we define

$$\operatorname{rank}(f) := \min \left\{ r \in \mathbb{N} : f = \sum_{p=1}^{r} \varphi_{1p} \otimes \cdots \otimes \varphi_{dp} \right\} \tag{10}$$

provided $f \neq 0$. The zero function is defined to have rank 0 and we say $\operatorname{rank}(f) = \infty$ if such a decomposition is not possible.

We will call a function f with $\operatorname{rank}(f) \leq r$ a *rank-r function*. Such a function may be written as a sum of r separable functions but possibly fewer. A decomposition of the form

$$f = \sum_{p=1}^{r} \varphi_{1p} \otimes \cdots \otimes \varphi_{dp} \tag{11}$$

[1]Point values of L^p-functions are undefined in general. So equations like these are taken to implicitly mean *almost everywhere*. Anyway all functions that arise in our applications will at least be continuous and so this is really not a point of great concern.

[2]We clarify our terminologies: A multipartite function is one for which the arguments $\mathbf{x}_1, \ldots, \mathbf{x}_d$ can come from any $X_1, \ldots, X_d$ but a multivariate function, in the usual sense of the word, is one where $X_1, \ldots, X_d$ are (measurable) subsets of $\mathbb{R}$. For example, while

$$g(u, v, w, x, y, z) = \varphi_1(u, v)\varphi_2(w)\varphi_3(x, y, z)$$

is not separable in the multivariate sense, it is separable in the multipartite sense: for $\mathbf{x}_1 = (u, v)$, $\mathbf{x}_2 = w$, $\mathbf{x}_3 = (x, y, z)$,

$$g(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \varphi_1(\mathbf{x}_1)\varphi_2(\mathbf{x}_2)\varphi_3(\mathbf{x}_3).$$

will be called a *rank-r multilinear decomposition*. Note that the qualificative 'rank-r' will always mean 'rank not more than r'. If we wish to refer to a function f with rank exactly r, we will just specify that $\operatorname{rank}(f) = r$. In this case, the rank-r multilinear decomposition in (11) is of mininum length and we call it a *rank-retaining multilinear decomposition* of f.

A rank-1 function is both non-zero and decomposable, i.e., of the form $\varphi_1 \otimes \cdots \otimes \varphi_d$ where $\varphi_k \in L^2(X_k)$. This agrees precisely with the notion of a separable function. Observe that the inner product (and therefore the norm) on $L^2(X_1 \times \cdots \times X_d)$ of a rank-1 function splits into a product

$$\langle \varphi_1 \otimes \cdots \otimes \varphi_d, \psi_1 \otimes \cdots \otimes \psi_d \rangle = \langle \varphi_1, \psi_1 \rangle_1 \cdots \langle \varphi_d, \psi_d \rangle_d \quad (12)$$

where $\langle \cdot, \cdot \rangle_p$ denotes the inner product of $L^2(X_p)$. This inner product extends linearly to finite-rank elements of $L^2(X_1 \times \cdots \times X_d)$: for $f = \sum_{p=1}^{r} \varphi_{1p} \otimes \cdots \otimes \varphi_{dp}$ and $g = \sum_{q=1}^{s} \psi_{1q} \otimes \cdots \otimes \psi_{dq}$, we have

$$\langle f, g \rangle = \sum_{p,q=1}^{r,s} \langle \varphi_{1p}, \psi_{1q} \rangle_1 \cdots \langle \varphi_{dp}, \psi_{dq} \rangle_d.$$

In fact this is how a tensor product of Hilbert spaces (the right hand side of (8)) is usually defined, namely, as the completion of the set of finite-rank elements of $L^2(X_1 \times \cdots \times X_d)$ under this inner product.

When $X_1, \ldots, X_d$ are finite sets, then all functions in $L^2(X_1 \times \cdots \times X_d)$ are of finite rank (and may in fact be viewed as hypermatrices or tensors as discussed in Section II). Otherwise there will be functions in $L^2(X_1 \times \cdots \times X_d)$ of infinite rank. However, since we have assumed that $X_1, \ldots, X_d$ are σ-finite measurable spaces, the set of all finite-rank f will always be dense in $L^2(X_1 \times \cdots \times X_d)$ by the Stone-Weierstrass theorem.

The next statement is a straightforward observation about multilinear decompositions of finite-rank functions but since it is central to this article we state it as a theorem. It is also tempting to call the decomposition a 'singular value decomposition' given its similarities with the usual matrix singular value decomposition (cf. Example 4).

Theorem 2 ('Singular value decomposition' for multipartite functions). Let $f \in L^2(X_1 \times \cdots \times X_d)$ be of finite rank. Then there exists a rank-r multilinear decomposition

$$f = \sum_{p=1}^{r} \sigma_p \varphi_{1p} \otimes \cdots \otimes \varphi_{dp} \quad (13)$$

such that

$$r = \operatorname{rank}(f), \quad (14)$$

the functions $\varphi_{kp} \in L^2(X_p)$ are of unit norm,

$$\|\varphi_{kp}\| = 1 \quad \text{for all } k = 1, \ldots, d, \quad p = 1, \ldots, r, \quad (15)$$

the coefficients $\sigma_1, \ldots, \sigma_r$ are real positive, and

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0. \quad (16)$$

Proof: This requires nothing more than rewriting the sum in (11) as a linear combination with the positive σ_p's accounting for the norms of the summands and then re-indexing them in descending order of magnitudes. ∎

While the usual singular value decomposition of a matrix would also have properties (14), (15), and (16), the one crucial difference here is that our 'singular vectors' $\varphi_{k1}, \ldots, \varphi_{kr}$ in (13) will only be of unit norm but will not in general be orthonormal. Given this, we will not expect properties like the Eckhart-Young theorem, or that $\sigma_1^2 + \cdots + \sigma_r^2 = \|f\|^2$, etc, to hold for (13) (cf. Section VI for more details).

One may think of the multilinear decomposition (13) as being similar in spirit, although not in substance, to Kolmogorov's superposition principle [40]; the main message of which is that:

"*There are no true multivariate functions.*"

More precisely, the principle states that continuous functions in multiple variables can be expressed as a composition of a univariate function with other univariate functions. For readers not familiar with this remarkable result, we state here a version of it due to Kahane [39].

Theorem 3 (Kolmogorov superposition). Let $f : [0, 1]^d \to \mathbb{R}$ be continuous. Then there exist constants $\lambda_1, \ldots, \lambda_d \in \mathbb{R}$ and Lipschitz continuous functions $\varphi_1, \ldots, \varphi_d : [0, 1] \to [0, 1]$ such that

$$f(x_1, \ldots, x_d) = \sum_{p=1}^{2d+1} g(\lambda_1 \varphi_p(x_1) + \cdots + \lambda_d \varphi_p(x_d)).$$

It is in general not easy to determine g and $\varphi_1, \ldots, \varphi_{2d+1}$ given such a function f. A multilinear decomposition of the form (13) alleviates this by allowing g to be the simplest multivariate function, namely, the product function,

$$g(t_1, \ldots, t_d) = t_1 t_2 \cdots t_d, \quad (17)$$

and unlike the univariate g in Theorem III, the g in (17) works universally for any function f — only the φ_p's need to be constructed. Furthermore, (13) applies more generally to functions on a product of general domains $X_1, \ldots, X_d$ whereas Theorem 2 only applies if they are compact intervals of $\mathbb{R}$.

At this stage, it would be instructive to give a few examples for concreteness.

Example 4 (Singular value decomposition). Let $A \in \mathbb{C}^{m \times n}$ be a matrix of rank r. Then it can be decomposed in infinitely many ways into a sum of rank-1 terms as

$$A = \sum_{p=1}^{r} \sigma_p \mathbf{u}_p \mathbf{v}_p^* \quad (18)$$

where $\mathbf{u}_p \in \mathbb{C}^m$ and $\mathbf{v}_p \in \mathbb{C}^n$ are unit-norm vectors and $\sigma_1 \geq \cdots \geq \sigma_r > 0$. Note that if we regard A as a complex-valued function on its row and column indices i and j as described earlier in Section II, then (18) may be written as

$$a(i, j) = \sum_{p=1}^{r} \sigma_p u_p(i) v_p(j),$$

which clearly is the same as (9). The singular value decomposition (SVD) of A yields one such decomposition, where $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ and $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ are both orthonormal. But in general a rank-retaining decomposition of the form (13) will not have such a property.

Example 5 (Schmidt decomposition). The previous example can be generalized to infinite dimensions. Let $A : \mathbb{H}_1 \to \mathbb{H}_2$ be a compact operator (also known as a completely continuous operator) between two separable Hilbert spaces. Then the Schmidt decomposition theorem says that there exist orthonormal basis $\{\varphi_p \in \mathbb{H}_2 : p \in \mathbb{N}\}$ and $\{\psi_p \in \mathbb{H}_1 : p \in \mathbb{N}\}$ so that

$$Af = \sum_{p=1}^{\infty} \sigma_p \langle \psi_p, f \rangle \varphi_p \tag{19}$$

for every $f \in \mathbb{H}_1$. In tensor product notation, (19) becomes

$$A = \sum_{p=1}^{\infty} \sigma_p \varphi_p \otimes \psi_p^*.$$

where ψ_p^* denotes the dual form of ψ_p.

Examples 4 and 5 are well-known but they are bipartite examples, i.e. $d = 2$ in (13). This article is primarily concerned with the d-partite case where $d \geq 3$, which has received far less attention. As we have alluded to in the previous section, the identification techniques in this article will rely crucially on the fact that $d \geq 3$.

Example 6. Let $A \in \mathbb{C}^{l \times m \times n}$ be a 3-dimensional hypermatrix. The outer product of three vectors $\mathbf{u} \in \mathbb{C}^l$, $\mathbf{v} \in \mathbb{C}^m$, $\mathbf{w} \in \mathbb{C}^n$ is defined by

$$\mathbf{u} \otimes \mathbf{v} \otimes \mathbf{w} = (u_i v_j w_k)_{i,j,k=1}^{l,m,n} \in \mathbb{C}^{l \times m \times n}.$$

The rank of A is defined to be the minimum $r \in \mathbb{N}$ such that A can be written in the form

$$A = \sum_{p=1}^{r} \sigma_p \mathbf{u}_p \otimes \mathbf{v}_p \otimes \mathbf{w}_p, \tag{20}$$

and if $A = 0$, then its rank is set to be 0. This agrees of course with our use of the word rank in (10), the only difference is notational, since (20) may be written in the form

$$a(i, j, k) = \sum_{p=1}^{r} \sigma_p u_p(i) v_p(j) w_p(k).$$

This definition of rank is invariant under the natural action[3] $\mathrm{GL}_l(\mathbb{C}) \times \mathrm{GL}_m(\mathbb{C}) \times \mathrm{GL}_n(\mathbb{C})$ on $\mathbb{C}^{l \times m \times n}$ [21, Lemma 2.3], i.e., for any $X \in \mathrm{GL}_l(\mathbb{C})$, $Y \in \mathrm{GL}_m(\mathbb{C})$, $Z \in \mathrm{GL}_n(\mathbb{C})$,

$$\operatorname{rank}((X, Y, Z) \cdot A) = \operatorname{rank}(A). \tag{21}$$

The definition also extends easily to d-dimensional hypermatrices in $\mathbb{C}^{n_1 \times \cdots \times n_d}$ and when $d = 2$ reduces to the usual definition in Example III for matrix rank. This definition is due to F. L. Hitchcock [37] and is often called *tensor rank*. The only difference here is that our observation in Theorem 2 allows us to impose the conditions

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r$$

and

$$\|\mathbf{u}_p\| = \|\mathbf{v}_p\| = \|\mathbf{w}_p\| = 1, \quad p = 1, \ldots, r, \tag{22}$$

[3] $\mathrm{GL}_n(\mathbb{C}) := \{A \in \mathbb{C}^{n \times n} : \det(A) \neq 0\}$ denotes the general linear goup of nonsingular $n \times n$ complex matrices.

while leaving $\operatorname{rank}(A)$ unchanged, thus bringing (20) closer in form to its matrix cousin (18). What is lost here is that the sets $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}, \{\mathbf{v}_1, \ldots, \mathbf{v}_r\}, \{\mathbf{w}_1, \ldots, \mathbf{w}_r\}$ can no longer be chosen to be orthonormal as in Example 4, the unit norm condition (22) is as far as we may go. In fact for a generic $A \in \mathbb{C}^{l \times m \times n}$, we will always have

$$r > \max(l, m, n),$$

and $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}, \{\mathbf{v}_1, \ldots, \mathbf{v}_r\}, \{\mathbf{w}_1, \ldots, \mathbf{w}_r\}$ will be overcomplete sets in $\mathbb{C}^l, \mathbb{C}^m, \mathbb{C}^n$ respectively.

Perhaps it is worthwhile saying a word concerning our use of the words 'tensor' and 'hypermatrix': A d-tensor or order-d tensor is an element of a tensor product of d vector spaces $\mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_d$; a d-dimensional hypermatrix is an element of $\mathbb{C}^{n_1 \times \cdots \times n_d}$. If we choose bases on $\mathbb{V}_1, \ldots, \mathbb{V}_d$, then any d-tensor $\mathbf{A} \in \mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_d$ will have a unique coordinate representation as a d-dimensional hypermatrix $A \in \mathbb{C}^{n_1 \times \cdots \times n_d}$, where $n_k = \dim(\mathbb{V}_k)$. A notion defined on a hypermatrix is only defined on the tensor (that is represented in coordinates by the hypermatrix) if that notion is independent of the choice of bases. So the use of the word 'tensor rank' is in fact well justified because of (21). For more details, we refer the reader to [47].

IV. Uniqueness of Multilinear Decompositions

In Theorem 2, we chose the coefficients to be in descending order of magnitude and require the factors in each separable term to be of unit norm. This is largely to ensure as much uniqueness in the multilinear decomposition as generally possible. However there remain two obvious ways to obtain trivially different multilinear decompositions: (i) one may scale the factors $\varphi_{1p}, \ldots, \varphi_{dp}$ by arbitrary unimodulus complex numbers as long as their product is 1; (ii) when two or more successive coefficients are equal, their orders in the sum may be arbitrarily permuted. We will call a multilinear decomposition of f that meets the conditions in Theorem 2 *essentially unique* if the only other such decompositions of f differ in one or both of these manners.

It is perhaps astonishing that when $d > 2$, a sufficient condition for essential uniqueness can be derived with relatively mild conditions on the factors. This relies on the notion of Kruskal rank, which we will now define.

Definition 7. Let $\Phi = \{\varphi_1, \ldots, \varphi_r\}$ be a finite collection of vectors of unit norm in $L^2(X_1 \times \cdots \times X_d)$. The **Kruskal rank** of Φ, denoted $\operatorname{krank} \Phi$, is the largest $k \in \mathbb{N}$ so that every k-element subset of Φ contains linearly independent elements.

This notion was originally introduced in [41]. It is related to the notion of *spark* introduced in compressed sensing [27], [33], defined as the smallest $k \in \mathbb{N}$ so that there is at least one k-element subset of Φ that is linearly dependent. The relation is simple to describe, $\operatorname{spark} \Phi = \operatorname{krank} \Phi + 1$, and it follows immediately from the respective definitions. It is clear that $\dim \operatorname{span} \Phi \geq \operatorname{krank} \Phi$.

We now generalize Kruskal's famous result [41], [61] to tensor products of arbitrary Hilbert spaces, possibly of infinite dimensions. But first let us be more specific about essential uniqueness.

Definition 8. We shall say that a multilinear decomposition of the form (13) (satisfying both (16) and (15)) is **essentially unique** if given another such decomposition,

$$\sum_{p=1}^{r} \sigma_p \varphi_{1p} \otimes \cdots \otimes \varphi_{dp} = f = \sum_{p=1}^{r} \lambda_p \psi_{1p} \otimes \cdots \otimes \psi_{dp},$$

we must have (i) the coefficients $\sigma_p = \lambda_p$ for all $p = 1, \ldots, r$; and (ii) the factors $\varphi_{1p}, \ldots, \varphi_{dp}$ and $\psi_{1p}, \ldots, \psi_{dp}$ differ at most via unimodulus scaling, i.e.

$$\varphi_{1p} = e^{i\theta_{1p}} \psi_{1p}, \ldots, \varphi_{dp} = e^{i\theta_{dp}} \psi_{dp} \tag{23}$$

where $\theta_{1p} + \cdots + \theta_{dp} \equiv 0 \bmod 2\pi$, for all $p = 1, \ldots, r$. In the event when successive coefficients are equal, $\sigma_{p-1} > \sigma_p = \sigma_{p+1} = \cdots = \sigma_{p+q} > \sigma_{p+q+1}$, the uniqueness of the factors in (ii) is only up to relabelling of indices, i.e. $p, \ldots, p+q$.

Lemma 9 (Infinite-dimensional Kruskal uniqueness). Let $f \in L^2(X_1 \times \cdots \times X_d)$ be of finite rank. Then a multilinear decomposition of the form

$$f = \sum_{p=1}^{r} \sigma_p \varphi_{1p} \otimes \cdots \otimes \varphi_{dp} \tag{24}$$

is both essentially unique and rank-retaining, i.e., $r = \operatorname{rank} f$, if the following condition is satisfied:

$$2r + d - 1 \leq \sum_{k=1}^{d} \operatorname{krank} \Phi_k, \tag{25}$$

where $\Phi_k = \{\varphi_{k1}, \ldots, \varphi_{kr}\}$ for $k = 1, \ldots, d$.

Proof: Consider the subspaces $\mathbb{V}_k = \operatorname{span}(\varphi_{k1}, \ldots, \varphi_{kr})$ in $L^2(X_k)$ for each $k = 1, \ldots, d$. Observe that $f \in \mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_d$. Clearly $\dim(\mathbb{V}_k) \leq r$ and so $\dim(\mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_d) \leq r^d$. Now, if we could apply Kruskal's uniqueness theorem [41] to the finite-dimensional space $\mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_d$, then we may immediately deduce both the uniqueness and rank-retaining property of (24). However there is one caveat: We need to show that Kruskal rank does not change under restriction to a subspace, i.e. the value of $\operatorname{krank}\{\varphi_{k1}, \ldots, \varphi_{kr}\}$ in (25) is the same whether we regard $\varphi_{k1}, \ldots, \varphi_{kr}$ as elements of $L^2(X_k)$ or as elements of the subspace $\mathbb{V}_k$. But this just follows from the simple fact that linear independence has precisely this property, i.e., if $v_1, \ldots, v_n \in \mathbb{U} \subseteq \mathbb{V}$ are linearly independent in the vector space $\mathbb{V}$, then they are linearly independent in the subspace $\mathbb{U}$. ∎

It follows immediately why we usually need $d \geq 3$ for identifiability.

Corollary 10. A necessary condition for Kruskal's inequality (25) to hold is that $d \geq 3$.

Proof: If $d = 2$, then $2r + d - 1 = 2r + 1 > \operatorname{krank} \Phi_1 + \operatorname{krank} \Phi_2$ since the Kruskal rank of of r vectors cannot exceed r. Likewise for $d = 1$. ∎

Lemma 9 shows that the condition in (25) is sufficient to ensure uniqueness and it is known that the condition is not necessary. In an appropriate sense, the condition is sharp [24]. We should note that the version of Lemma 9 that we state here for general $d \geq 3$ is due to Sidiropoulos and Bro [61]. Kruskal's original version [41] is only for $d = 3$.

The main problem with Lemma 9 is that the condition (25) is difficult to check since the right-hand side cannot be readily computed. See Section VIII-F for a discussion.

Kruskal's result also does not tell us how often multilinear decompositions are unique. In the event when the sets $X_1, \ldots, X_d$ are finite, $L^2(X_1 \times \cdots \times X_d) \cong \mathbb{C}^{n_1 \times \cdots \times n_d}$ where $n_1 = \#X_1, \ldots, n_d = \#X_d$, and there is a simple result on uniqueness based simply on a dimension count. Note that the dimension of $L^2(X_1 \times \cdots \times X_d)$ is the product $n_1 \cdots n_d$ and the number of parameters needed to describe a separable element of the form $\lambda \varphi_1 \otimes \cdots \otimes \varphi_d$ where $\varphi_1, \ldots, \varphi_d$ are of unit norm is $n_1 + \cdots + n_d - d + 1$ (each φ_k requires $n_k - 1$ parameters because of the unit norm constraint, the last '+1' accounts for the coefficient λ). We call the number

$$\left\lceil \frac{\prod_{k=1}^{d} n_k}{1 - d + \sum_{k=1}^{d} n_k} \right\rceil$$

the *expected rank* of $L^2(X_1 \times \cdots \times X_d)$, since it is heuristically the minimum r expected for a multilinear decomposition (13).

Proposition 11. Let the notations be as above. If $f \in L^2(X_1 \times \cdots \times X_d)$ has rank smaller than the expected rank, i.e.

$$\operatorname{rank}(f) < \left\lceil \frac{\prod_{k=1}^{d} n_k}{1 - d + \sum_{k=1}^{d} n_k} \right\rceil,$$

then f admits at most a finite number of distinct rank-retaining decompositions.

This proposition has been proved in several cases, including symmetric tensors [14], but the proof still remains incomplete for tensors of most general form [1], [12].

V. Estimation of Multilinear Decompositions

In practice we would only have at our disposal $\hat{f}$, a measurement of f corrupted by noise. Recall that our model for f takes the form

$$f(\mathbf{x}_1, \ldots, \mathbf{x}_d) = \sum_{p=1}^{r} \alpha_p \prod_{k=1}^{d} \varphi_{kp}(\mathbf{x}_k). \tag{26}$$

Then we would often have to solve an approximation problem corresponding to (26) of the form

$$\operatorname*{argmin}_{\boldsymbol{\alpha} \in \mathbb{C}^r,\ \|\varphi_{kp}\| = 1} \left\| \hat{f} - \sum_{p=1}^{r} \alpha_p \prod_{k=1}^{d} \varphi_{kp} \right\|, \tag{27}$$

which we will call a *best rank-r approximation problem*. A solution to (27), if exists, will be called a best rank-r approximation of $\hat{f}$.

We will give some motivations as to why such an approximation is reasonable. Assuming that the norm in (27) is the L^2-norm and that the factors φ_{kp}, $p = 1, \ldots r$ and $k = 1, \ldots d$, have been determined in advance and we are just trying to estimate the parameters $\alpha_1, \ldots, \alpha_r$ from $\hat{f}^{(1)}, \ldots, \hat{f}^{(N)}$ a finite sample of size N of measurements of f corrupted by noise, then the solution of the approximation problem in (27) is in fact (i) a maximum likelihood estimator (MLE) if the noise is zero mean Gaussian, and (ii) a best linear unbiased estimator (BLUE) if the noise has zero mean and finite variance.

Of course in our identification problems, the factors φ_{kp}'s are not known and have to be estimated too. A probabilistic model in this situation would take us too far afield. Note that even for the case $d = 2$ and where the domain of f $X_1 \times X_2$ is a finite set, a case that essentially reduces to principal components analysis (PCA), a probabilistic model along the lines of [71] requires several strong assumptions and was only developed as late as 1999. The lack of a formal probabilistic model has not stopped PCA, proposed in 1901 [58], to be an invaluable tool in the intervening century.

VI. Existence of Best Multilinear Approximation

As we mentioned in the previous section, in realistic situation where measurements are corrupted by additive noise, one has to extract the factors φ_{kp}'s and α_p through solving an approximation problem (27), that we now write in a slightly different (but equivalent) form,

$$\operatorname*{argmin}_{\boldsymbol{\alpha}\in[0,\infty)^r,\ \|\varphi_{kp}\|=1} \left\| \hat{f} - \sum_{p=1}^{r} \alpha_p \prod_{k=1}^{d} \varphi_{kp} \right\|. \tag{28}$$

Note that by Theorem 2, we may assume that the coefficients $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_r)$ are real and nonnegative valued without any loss of generality. Such a form is also natural in applications given that α_p usually captures the magnitude of whatever quantity that is represented by the p summand.

We will see this problem, whether in the form (27) or (28), has no solution in general. We will first observe a somewhat unusual phenomenon in multilinear decomposition of d-partite functions where $d \geq 3$, namely, a sequence of rank-r functions (each with an rank-r multilinear decomposition) can converge to a limit that is not rank-r (has no rank-r multilinear decomposition).

Example 12 (Multilinear approximation of functions) For linearly independent $\varphi_1, \psi_1 : X_1 \to \mathbb{C}$, $\varphi_2, \psi_2 : X_2 \to \mathbb{C}$, $\varphi_3, \psi_3 : X_3 \to \mathbb{C}$, let $\hat{f} : X_1 \times X_2 \times X_3 \to \mathbb{C}$ be

$$\begin{aligned}\hat{f}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) :=\ & \psi_1(\mathbf{x}_1)\varphi_2(\mathbf{x}_2)\varphi_3(\mathbf{x}_3) \\ & + \varphi_1(\mathbf{x}_1)\psi_2(\mathbf{x}_2)\varphi_3(\mathbf{x}_3) + \varphi_1(\mathbf{x}_1)\varphi_2(\mathbf{x}_2)\psi_3(\mathbf{x}_3).\end{aligned}$$

For $n \in \mathbb{N}$, define

$$\begin{aligned}f_n(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) := n &\left[\varphi_1(\mathbf{x}_1) + \frac{1}{n}\psi_1(\mathbf{x}_1)\right]\left[\varphi_2(\mathbf{x}_2) + \frac{1}{n}\psi_2(\mathbf{x}_2)\right] \\ &\left[\varphi_3(\mathbf{x}_3) + \frac{1}{n}\psi_3(\mathbf{x}_3)\right] - n\varphi_1(\mathbf{x}_1)\varphi_2(\mathbf{x}_2)\varphi_3(\mathbf{x}_3).\end{aligned}$$

Then

$$\begin{aligned}f_n(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) - \hat{f}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \frac{1}{n}\big[&\psi_1(\mathbf{x}_1)\psi_2(\mathbf{x}_2)\varphi_3(\mathbf{x}_3) \\ &+ \psi_1(\mathbf{x}_1)\varphi_2(\mathbf{x}_2)\psi_3 \\ &+ \varphi_1(\mathbf{x}_1)\psi_2(\mathbf{x}_2)\psi_3(\mathbf{x}_3)\big] \\ &+ \frac{1}{n^2}\psi_1(\mathbf{x}_1)\psi_2(\mathbf{x}_2)\psi_3(\mathbf{x}_3).\end{aligned}$$

Hence

$$\|\hat{f} - f_n\| = O\left(\frac{1}{n}\right). \tag{29}$$

Lemma 13. In Example VI, $\operatorname{rank}(\hat{f}) = 3$ iff φ_i, ψ_i are linearly independent, $i = 1, 2, 3$. Furthermore, it is clear that $\operatorname{rank}(f_n) \leq 2$ and

$$\lim_{n\to\infty} f_n = \hat{f}.$$

Note that our fundamental approximation problem may be regarded as the approximation problem

$$\operatorname{argmin}\{\|\hat{f} - f\| : \operatorname{rank}(f) \leq r\}, \tag{30}$$

followed by a decomposition problem

$$f = \sum_{p=1}^{r} \alpha_p \prod_{k=1}^{d} \varphi_{kp},$$

which always exists for an f with $\operatorname{rank}(f) \leq r$. The discussion above shows that there are target functions $\hat{f}$ for which

$$\operatorname{argmin}\{\|\hat{f} - f\| : \operatorname{rank}(f) \leq r\} = \varnothing,$$

and thus (28) or (30) does not need to have a solution in general. This is such a crucial point that we are obliged to formally state it.

Theorem 14. For $d \geq 3$, the best approximation of a d-partite function by a sum of p products of d separable functions does not exist in general.

Proof: Take the tripartite function $\hat{f} \in L^2(X_1 \times X_2 \times X_3)$ in Example VI. Suppose we seek a best rank-2 approximation, in other words, we seek to solve the minimization problem

$$\operatorname*{argmin}_{\|g_k\|=\|h_k\|=1,\ \gamma,\eta\geq 0} \|\hat{f} - \gamma\, g_1 \otimes g_2 \otimes g_3 - \eta h_1 \otimes h_2 \otimes h_3\|.$$

Now, the *infimum*,

$$\inf_{\|g_k\|=\|h_k\|=1,\ \gamma,\eta\geq 0} \|\hat{f} - \gamma\, g_1 \otimes g_2 \otimes g_3 - \eta h_1 \otimes h_2 \otimes h_3\| = 0$$

since we may choose $n \in \mathbb{N}$ sufficiently large,

$$g_k = \frac{\varphi_k + n^{-1}\psi_k}{\|\varphi_k + n^{-1}\psi_k\|}, \quad h_k = \frac{\varphi_k}{\|\varphi_k\|},$$

for $k = 1, 2, 3$,

$$\begin{aligned}\gamma &= n\|\varphi_1 + n^{-1}\psi_1\|\|\varphi_2 + n^{-1}\psi_2\|\|\varphi_3 + n^{-1}\psi_3\|, \\ \eta &= n\|\varphi_1\|\|\varphi_2\|\|\varphi_3\|,\end{aligned}$$

so as make $\|\hat{f} - \gamma\, g_1 \otimes g_2 \otimes g_3 - \eta h_1 \otimes h_2 \otimes h_3\|$ as small as we desired by virtue of (29). However there is no rank-2 function $\gamma\, g_1 \otimes g_2 \otimes g_3 + \eta h_1 \otimes h_2 \otimes h_3$ for which

$$\|\hat{f} - \gamma\, g_1 \otimes g_2 \otimes g_3 - \eta h_1 \otimes h_2 \otimes h_3\| = 0.$$

In other words, the zero infimum can never be attained. ∎

Our construction above is based on an earlier construction in [21]. The first such example was given in [4], which also contains the very first definition of border rank. We will define it here for d-partite functions. When $X_1, \dots, X_d$ are finite sets, this reduces to the original definition in [4] for hypermatrices.

Definition 15. Let $f \in L^2(X_1 \times \cdots \times X_d)$. The **border rank** of f is defined as

$$\begin{aligned}\overline{\operatorname{rank}}(f) = \min\{r \in \mathbb{N} : \inf\|f - g\| = 0 \\ \text{over all } g \text{ with } \operatorname{rank}(g) = r\}.\end{aligned}$$

We say $\overline{\operatorname{rank}}(f) = \infty$ if such a finite r does not exist.

Clearly we would always have that

$$\overline{\operatorname{rank}}(f) \leq \operatorname{rank}(f).$$

The discussions above show that strict inequality can occur. In fact, for the $\hat{f}$ in Example VI, $\overline{\operatorname{rank}}(\hat{f}) = 2$ while $\operatorname{rank}(\hat{f}) = 3$.

We would like to mention here that this problem applies to operators too. Optimal approximation of an operator by a sum of tensor/Kronecker products of lower-dimensional operators, which arises in numerical operator calculus [3], is in general an ill-posed problem whose existence cannot be guaranteed.

Example 16. (Multilinear approximation of operators). For linearly independent operators $\Phi_i, \Psi_i : V_i \to W_i$, $i = 1, 2, 3$, let $\widehat{T} : V_1 \otimes V_2 \otimes V_3 \to W_1 \otimes W_2 \otimes W_3$ be

$$\widehat{T} := \Psi_1 \otimes \Phi_2 \otimes \Phi_3 + \Phi_1 \otimes \Psi_2 \otimes \Phi_3 + \Phi_1 \otimes \Phi_2 \otimes \Psi_3. \quad (31)$$

If Φ_i, Ψ_i's are all finite-dimensional and represented in coordinates as matrices, then '$\otimes$' may be taken to be Kronecker product of matrices. For $n \in \mathbb{N}$,

$$T_n := n\left[\Phi_1 + \frac{1}{n}\Psi_1\right] \otimes \left[\Phi_2 + \frac{1}{n}\Psi_2\right] \otimes \left[\Phi_3 + \frac{1}{n}\Psi_3\right] - n\Phi_1 \otimes \Phi_2 \otimes \Phi_3.$$

Then

$$\lim_{n\to\infty} T_n = \widehat{T}.$$

An example of an operator that has the form in (31) is the $3m$-dimensional Laplacian Δ_{3m}, which can be expressed in terms of the m-dimensional Laplacian Δ_m as

$$\Delta_{3m} = \Delta_m \otimes I \otimes I + I \otimes \Delta_m \otimes I + I \otimes I \otimes \Delta_m.$$

There are several simple but artificial ways to alleviate the issue of non-existent best approximant. Observe from the proof of Theorem VI that the coefficients in the approximant γ, η becomes unbounded in the limit. Likewise we see this happening in Example VI. In fact this must *always* happen — in the event when a function or operator is approximated by a rank-r function, i.e.

$$\left\| \hat{f} - \sum_{p=1}^{r} \alpha_p \prod_{k=1}^{d} \varphi_{kp} \right\| \quad \text{or} \quad \left\| \widehat{T} - \sum_{p=1}^{r} \alpha_p \bigotimes_{k=1}^{d} \Phi_{kp} \right\|, \quad (32)$$

and if a best approximation does not exist, then the r coefficients $\alpha_1, \ldots, \alpha_r$ must *all* diverge in magnitude to $+\infty$ as the approximant converges to the infimum of the norm loss function in (32). This result was first established in [21, Proposition 4.9].

So a simple but artificial way to prevent the nonexistence issue is to simply limit the sizes of the coefficients $\alpha_1, \ldots, \alpha_r$ in the approximant. One way to achieve this is regularization [48], [57] — adding a regularization term to our objective function in (28) to penalize large coefficients. A common choice is Tychonoff regularization, which uses a sum-of-squares regularization term:

$$\operatorname*{argmin}_{\boldsymbol{\alpha}\in[0,\infty)^r,\ \|\varphi_{kp}\|=1} \left\| \hat{f} - \sum_{p=1}^{r} \alpha_p \prod_{k=1}^{d} \varphi_{kp} \right\| + \lambda \sum_{p=1}^{r} |\alpha_p|^2. \quad (33)$$

Here, λ is an arbitrarily chosen regularization parameter. It can be seen that this is equivalent to constraining the sizes $\alpha_1, \ldots, \alpha_r$ to $\sum_{p=1}^{r} |\alpha_p|^2 = \rho$, with ρ being determined a posteriori from λ. The main drawback of such constraints is that ρ and λ are arbitrary, and that they generally have no physical meaning.

More generally, one may alleviate the nonexistence issue by restricting the optimization problem (30) to a compact subset of its non-compact feasible set

$$\{f \in L^2(X_1 \times \cdots \times X_d) : \operatorname{rank}(f) \leq r\}.$$

Limiting the sizes of $\alpha_1, \ldots, \alpha_r$ is a special case but there are several other simple (but also artificial) strategies. In [17], the factors $\varphi_{k1}, \ldots, \varphi_{kp}$ are required to be orthogonal *for all* $k \in \{1, \ldots, d\}$, i.e.

$$\langle \varphi_{kp}, \varphi_{kq} \rangle_k = \delta_{pq}, \quad p, q = 1, \ldots, r, \quad k = 1, \ldots, d. \quad (34)$$

This remedy is acceptable only in very restrictive conditions. In fact a necessary condition for this to work is that

$$r \leq \min_{k=1,\ldots,d} \dim L^2(X_k).$$

It is also trivial to see that imposing orthogonality between the separable factors removes this problem

$$\langle \varphi_{1p} \otimes \cdots \otimes \varphi_{dp}, \varphi_{1q} \otimes \cdots \otimes \varphi_{dq} \rangle = \delta_{pq}, \quad p, q = 1, \ldots, r. \quad (35)$$

This constraint is slightly less restrictive — by (12), it is equivalent to requiring (34) *for some* $k \in \{1, \ldots, d\}$. Both (34) and (35) are nonetheless so restrictive as to exclude the most useful circumstances for the model (13), which usually involves factors that have no reason to be orthogonal, as we will see in Section IX. In fact, Kruskal's uniqueness condition is such a potent tool precisely because it does not require orthogonality.

The conditions (34), (35), and (33) all limit the feasible sets for the original approximation (28) to a much smaller compact subset of the original feasible set. This is not the case for nonnegative constraints. In [48] it was shown that the following best rank-r approximation problem for a nonnegative-valued $\hat{f}$ and where the coefficients α_p and factors φ_{kp} of the approximants are also nonnegative valued, i.e.

$$\operatorname*{argmin}_{\boldsymbol{\alpha}\in[0,\infty)^r,\ \|\varphi_{kp}\|=1,\ \varphi_{kp}\geq 0} \left\| \hat{f} - \sum_{p=1}^{r} \alpha_p \prod_{k=1}^{d} \varphi_{kp} \right\|,$$

always has a solution. The feasible set in this case is non-compact and has nonempty interior within the feasible set of our original problem (28). The nonnegativity constraints are natural in some applications, such as the fluorescence spectroscopy one described in Section IX-F, where φ_{kp} represent intensities and concentrations, and are therefore nonnegative valued.

There are two major problems with imposing artificial constraints simply to force a solution: How do we know a priori that the solution that we seek would meet those constraints? But more importantly, perhaps the model is ill-posed and a solution indeed should not exist? To illustrate the case in point with a more commonplace example, suppose

we want to find a maximum likelihood estimator $X \in \mathbb{R}^{n\times n}$ for the covariance Σ of independent samples $\mathbf{y}_1, \dots, \mathbf{y}_m \sim N(0, \Sigma)$. This would lead us to a semi-definite programming problem

$$\operatorname*{argmin}_{X\succ 0} \operatorname{tr}(X^{-1}Y) - \log\det(X) \tag{36}$$

where $Y = \frac{1}{m}\sum_{i=1}^m \mathbf{y}_i\mathbf{y}_i^{\mathsf{T}}$. However the problem will not have a solution when the number of samples is smaller than the dimension, i.e., $m < n$, as the infimum of the loss function in (36) cannot be attained by any X in the feasible set. This is an indication that we should seek more samples (so that we could get $m \geq n$, which will guarantee the attainment of the infimum) or use a different model (e.g., determine if X^{-1} might have some a priori zero entries due to statistical independence of the variables). It is usually unwise to impose artificial constraints on the covariance matrix X just so that the loss function in (36) would attain an infimum on a smaller feasible set — the thereby obtained 'solution' may bear no relation to the true solution that we want.

Our goal in Section VIII-A is to define a type of physically meaningful constraints via the notion of *coherence*. It ensures the existence of a unique minimum, but not via an artificial limitation of the optimization problem to a convenient subset of the feasible set. In the applications we discuss in Section IX, we will see that it is natural to expect existence of a solution when coherence is small enough, but not otherwise. So when our model is ill-posed or ill-conditioned, we are warned by the size of the coherence and could seek other remedies (collect more measurements, use a different model, etc) instead of forcing a 'solution' that bears no relation to reality. But before we get to that we will examine why, unlike in compressed sensing and matrix completion, the approximation of rank by a ratio of nuclear and spectral norms could not be expected to work here.

VII. Nuclear and Spectral Norms

We introduce the notion of nuclear and spectral norms for multipartite functions. Our main purpose is to see if they could be used to alleviate the problem discussed in Section VI, namely, that a d-partite function may not have a best approximation by a sum of r separable functions.

The definition of nuclear norm follows naturally from the definition of rank in Section III.

Definition 17. We define the **nuclear norm** (or **Schatten 1-norm**) of $f \in L^2(X_1 \times \cdots \times X_d)$ as

$$\|f\|_* := \inf\left\{\left[\sum_{p=1}^{\infty}\lambda_p\right] : f = \sum_{p=1}^{\infty}\lambda_p\varphi_{1p}\otimes\cdots\otimes\varphi_{dp}, \quad \|\varphi_{kp}\| = 1,\ \lambda_p \geq \lambda_{p+1} > 0\right\}. \tag{37}$$

Note that for rank-1 functions, we always have that

$$\|\varphi_1\otimes\cdots\otimes\varphi_d\|_* = \|\varphi_1\|\cdots\|\varphi_d\|. \tag{38}$$

A finite rank function always has finite nuclear norm but in general a function in $L^2(X_1 \times \cdots \times X_d)$ need not have finite nuclear norm.

The definition of the spectral norm of a multipartite function is motivated by the fact the usual spectral norm of a matrix A equals the maximal absolute value of its inner product $\operatorname{tr}(AX)$ with rank-1 unit-norm matrices $X = \mathbf{u}\mathbf{v}^*$, $\|\mathbf{u}\|_2 = \|\mathbf{v}\|_2 = 1$.

Definition 18. We define the **spectral norm** of $f \in L^2(X_1 \times \cdots \times X_d)$ as

$$\|f\|_\sigma := \sup\big\{|\langle f, \varphi_1\otimes\cdots\otimes\varphi_d\rangle| : \|\varphi_1\| = \cdots = \|\varphi_d\| = 1\big\}. \tag{39}$$

Here we write $\|\cdot\|$ for the L^2-norms in $L^2(X_k)$, $k = 1, \dots, d$. Alternatively, we may also use $\operatorname{Re}\langle f, \varphi_1\otimes\cdots\otimes\varphi_d\rangle$ in place of $|\langle f, \varphi_1\otimes\cdots\otimes\varphi_d\rangle|$ in (39), which does not change its value. Note that a function in $L^2(X_1 \times \cdots \times X_d)$ always has finite spectral norm.

The fact that (37) and (39) define norms on $L^2(X_1 \times \cdots \times X_d)$ follows from the standard Minkowski gauge argument [22]. Suppose $X_1, \dots, X_d$ are finite sets of cardinalities $n_1, \dots, n_d \in \mathbb{N}$. The nuclear and spectral norms for the unipartite case ($d = 1$) are the ℓ^1- and ℓ^∞-norms for vectors in $\mathbb{C}^{n_1} = L^2(X_1)$. The nuclear and spectral norms for the bipartite case ($d = 2$) agrees with the usual nuclear and spectral norms for matrices in $\mathbb{C}^{n_1\times n_2} = L^2(X_1 \times X_2)$. For general $d \geq 3$, Definition VII yields a notion of nuclear norm[4] for hypermatrices in $\mathbb{C}^{n_1\times\cdots\times n_d} = L^2(X_1 \times \cdots \times X_d)$, while Definition VII agrees with the usual notion of spectral norm for hypermatrices [46].

Example 19 (Nuclear and spectral norms for 3-tensors). Let $T \in \mathbb{C}^{n_1\times n_2\times n_3}$. Then by Definition VII, we have

$$\|T\|_* = \inf\left\{\sum_{p=1}^{r}\lambda_p : T = \sum_{p=1}^{r}\lambda_p\mathbf{u}_p\otimes\mathbf{v}_p\otimes\mathbf{w}_p\right\},$$

where the infimum is taken over all linear combinations of complex vectors of unit 2-norm $\mathbf{u}_p \in \mathbb{C}^{n_1}$, $\mathbf{v}_p \in \mathbb{C}^{n_2}$, $\mathbf{w}_p \in \mathbb{C}^{n_3}$, with real positive coefficientss $\lambda_p \in [0, \infty)$, and $p = 1, \dots, r$, with $r \in \mathbb{N}$.

We shall write $\|\cdot\| = \|\cdot\|_2$. The spectral norm of T is

$$\begin{aligned}\|T\|_\sigma &= \sup\{|\langle T, \mathbf{u}\otimes\mathbf{v}\otimes\mathbf{w}\rangle| : \|\mathbf{u}\| = \|\mathbf{v}\| = \|\mathbf{w}\| = 1\}\\ &= \sup_{\mathbf{x},\mathbf{y},\mathbf{z}\neq\mathbf{0}}\frac{|T(\mathbf{x},\mathbf{y},\mathbf{z})|}{\|\mathbf{x}\|\|\mathbf{y}\|\|\mathbf{z}\|} = \|T\|_{2,2,2}.\end{aligned}$$

We have regarded T as a trilinear functional defined by $T(\mathbf{x},\mathbf{y},\mathbf{z}) = \sum_{i,j,k=1}^{n_1,n_2,n_3} t_{ijk}x_iy_jz_k$ and $\|T\|_{2,2,2}$ is its induced norm as defined in [46], [47]. Again, these clearly extend to any d-tensors. We will see in Lemma VII that the nuclear and spectral norms for tensors are dual to each other.

Note that we have used the term *tensors*, as opposed to hypermatrices, in the above example. In fact, Definition 17 defines nuclear norms for the tensors, not just their coordinate representations as hypermatrices (see our discussion after Example III), because of the following invariant properties.

Lemma 20. The nuclear and spectral norms for $\mathbb{C}^{n_1\times\cdots\times n_d}$ are unitarily invariant, i.e., invariant under the natural action

[4]The notion of a nuclear norm for tensors was originally introduced in Section 3 of our 2010 article (cf. http://arxiv.org/abs/1002.4935v1). However, it was ultimately removed in the published version [49] because of the page limit of the journal.

of $\mathrm{U}_{n_1}(\mathbb{C}) \times \cdots \times \mathrm{U}_{n_d}(\mathbb{C})$ where $\mathrm{U}_n(\mathbb{C})$ denotes the group of unitary matrices in $\mathbb{C}^{n\times n}$.

Proof: To avoid the clutter of indices, we will assume that $d = 3$. It is easy, although notationally cumbersome, to extend this to general $d \geq 3$. Let $(U, V, W) \in \mathrm{U}_{n_1}(\mathbb{C}) \times \mathrm{U}_{n_2}(\mathbb{C}) \times \mathrm{U}_{n_3}(\mathbb{C})$ and $T \in \mathbb{C}^{n_1\times n_2\times n_3}$. The natural action, given in coordinates by

$$(U, V, W) \cdot T = \left[\sum_{i,j,k=1}^{n_1,n_2,n_3} u_{ai} v_{bj} w_{ck} t_{ijk}\right]_{a,b,c=1}^{n_1,n_2,n_3},$$

has the property that if T has a multilinear decomposition of the form

$$T = \sum_{p=1}^{r} \lambda_p \mathbf{x}_p \otimes \mathbf{y}_p \otimes \mathbf{z}_p,$$

then

$$(U, V, W) \cdot T = \sum_{p=1}^{r} \lambda_p (U\mathbf{x}_p) \otimes (V\mathbf{y}_p) \otimes (W\mathbf{z}_p). \quad (40)$$

(40) is obvious when $r = 1$ and for general r follows from the linearity of the action, i.e., $(U, V, W) \cdot (S + T) = (U, V, W) \cdot S + (U, V, W) \cdot T$. We also need the simple fact that $\mathrm{U}_{n_1}(\mathbb{C}) \times \mathrm{U}_{n_2}(\mathbb{C}) \times \mathrm{U}_{n_3}(\mathbb{C})$ acts transitively on unit-norm rank-1 tensors, i.e., take any $\mathbf{x} \in \mathbb{C}^{n_1}$, $\mathbf{y} \in \mathbb{C}^{n_2}$, $\mathbf{z} \in \mathbb{C}^{n_3}$ of unit norm, then every other unit-norm rank-1 tensor may be written as $U\mathbf{x} \otimes V\mathbf{y} \otimes W\mathbf{z}$ for some $(U, V, W) \in \mathrm{U}_{n_1}(\mathbb{C}) \times \mathrm{U}_{n_2}(\mathbb{C}) \times \mathrm{U}_{n_3}(\mathbb{C})$. With these, it follows immediately from Definition 17 that nuclear norms satisfy

$$\|(U, V, W) \cdot T\|_* = \|T\|_*.$$

One may similarly show that the spectral norm is also unitarily invariant or deduce the fact from Lemma 21 below. ∎

Recall that on a Hilbert space $\mathbb{H}$ with inner product $\langle \cdot, \cdot \rangle$ the *dual norm* of a given norm $\|\cdot\|$ is defined as

$$\|f\|^\vee := \sup\{|\langle f, g\rangle| : \|g\| \leq 1\}.$$

If $\|\cdot\|$ is the norm induced by the inner product, then $\|\cdot\|^\vee = \|\cdot\|$; but in general they are different. Nevertheless one always have that $(\|\cdot\|^\vee)^\vee = \|\cdot\|$ and

$$|\langle f, g\rangle| \leq \|f\|^\vee \|g\| \quad (41)$$

for any $f, g \in \mathbb{H}$.

Since the ℓ^1- and ℓ^∞-norms on $\mathbb{C}^n$ are dual, as are the nuclear and spectral norms on $\mathbb{C}^{n_1\times n_2}$, one may wonder if it is true in general that the nuclear and spectral norms are dual to each other. This is in fact almost a tautology when $X_1, \ldots, X_d$ are finite.

Lemma 21. Let $X_1, \ldots, X_d$ be finite sets. Then nuclear and spectral norms on $L^2(X_1 \times \cdots \times X_d)$ satisfy

$$|\langle f, g\rangle| \leq \|f\|_\sigma \|g\|_* \quad (42)$$

and in fact

$$\|f\|_*^\vee = \|f\|_\sigma. \quad (43)$$

Proof: We first need to establish (42) without invoking (41). Since $X_1, \ldots, X_d$ are finite, any $g \in L^2(X_1 \times \cdots \times X_d)$ is of finite rank. Take any multilinear decomposition

$$g = \sum_{p=1}^{r} \lambda_p \varphi_{1p} \otimes \cdots \otimes \varphi_{dp}$$

where $r \in \mathbb{N}$ is arbitrary. Now

$$\begin{aligned}|\langle f, g\rangle| &\leq \sum_{p=1}^{r} |\lambda_p| |\langle f, \varphi_{1p} \otimes \cdots \otimes \varphi_{dp}\rangle| \\ &\leq \|f\|_\sigma \sum_{p=1}^{r} |\lambda_p|\end{aligned}$$

by definition of spectral norm. Taking infimum over all finite-rank decompositions, we arrive at (42) by definition of nuclear norm. Hence

$$\begin{aligned}\|f\|_*^\vee &= \sup\{|\langle f, g\rangle| : \|g\|_* \leq 1\} \\ &\leq \sup\{\|f\|_\sigma \|g\|_* : \|g\|_* \leq 1\} = \|f\|_\sigma.\end{aligned}$$

On the other hand, using (41) for $\|\cdot\|_*$ and $\|\cdot\|_*^\vee$, we get

$$\begin{aligned}\|f\|_\sigma &= \sup\{|\langle f, \varphi_1 \otimes \cdots \otimes \varphi_d\rangle| : \|\varphi_k\| = 1\} \\ &\leq \sup\{\|f\|_*^\vee \|\varphi_1 \otimes \cdots \otimes \varphi_d\|_* : \|\varphi_k\| = 1\} = \|f\|_*^\vee.\end{aligned}$$

where the last equality follows from (38). ∎

When $X_1, \ldots, X_d$ are only required to be σ-finite measurable spaces, we may use a limiting argument to show that (42) still holds for f of finite spectral norm and g of finite nuclear norm; a proper generalization of (43) is more subtle and we will leave this to future work since we do not require it in this article.

It is known [27] that the ℓ^1-norm is the *largest* convex underestimator[5] of the "ℓ^0-norm" on the ℓ^∞-norm unit ball [44] and that the nuclear norm is the *largest* convex underestimator of rank on spectral norm unit ball [30]. In particular,

$$\|\mathbf{x}\|_1 \leq \mathrm{nnz}(\mathbf{x}) \|\mathbf{x}\|_\infty$$

for all $\mathbf{x} \in \mathbb{C}^n$ while

$$\|X\|_* \leq \mathrm{rank}(X) \|X\|_\sigma$$

for all $X \in \mathbb{C}^{m\times n}$. The quantity $\mathrm{nnz}(\mathbf{x}) := \#\{i : x_i \neq 0\}$ is often called "ℓ^0-norm" even though it is not a norm (and neither a seminorm nor a quasinorm nor a pseudonorm).

We had suspected that the following generalization might perhaps be true, namely, rank, nuclear, and spectral norms as defined in (10), (37), and (39) would also satisfy the same inequality:

$$\|f\|_* \leq \mathrm{rank}(f) \times \|f\|_\sigma. \quad (44)$$

If true, this would immediately imply the same for border rank

$$\|f\|_* \leq \overline{\mathrm{rank}}(f) \times \|f\|_\sigma$$

by a limiting argument. Furthermore, (44) would provide a simple way to remedy the nonexistence problem highlighted in Theorem 14: One may use the ratio $\|f\|_*/\|f\|_\sigma$ as a 'proxy'

[5]Also called the *greatest convex minorant*, in this case also equivalent to the *Legendre-Frenchel biconjugate* or *convex biconjugate*.

in place of $\operatorname{rank}(f)$ and replace the condition $\operatorname{rank}(f) \leq r$ by the weaker condition $\|f\|_* \leq r\|f\|_\sigma$. The discussion in Section VI shows that there are $\hat{f}$ for which

$$\operatorname{argmin}\{\|\hat{f} - f\| : \operatorname{rank}(f) \leq r\} = \varnothing,$$

which really results from the fact that

$$\{f \in L^2(X_1 \times \cdots \times X_d) : \operatorname{rank}(f) \leq r\}$$

is not a closed set. But

$$\{f \in L^2(X_1 \times \cdots \times X_d) : \|f\|_* \leq r\|f\|_\sigma\} \tag{45}$$

is always closed (by the continuity of norms) and so for any $r \in \mathbb{N}$, the optimization problem

$$\operatorname{argmin}\{\|\hat{f} - f\| : \|f\|_* \leq r\|f\|_\sigma\}$$

always has a solution.

Unfortunately, (44) is not true when $d > 2$. The following example shows that nuclear norm is not an underestimator of rank on the spectral norm unit ball, and can in fact be arbitrarily larger than rank on the spectral norm unit ball.

Example 22 (Matrix multiplication). Let $T_n \in \mathbb{C}^{n^2 \times n^2 \times n^2}$ be the matrix multiplication tensor (cf. Applications 2 and 3 in [47, Section 15.3]). The well-known result of Strassen [67] implies that

$$\operatorname{rank}(T_n) \leq cn^{\log_2 7}$$

for some $c > 0$ and for n sufficiently large. On the other hand, Derksen [25] has recently established the exact values for the nuclear and spectral norm of T_n:

$$\|T_n\|_* = n^3, \quad \|T_n\|_\sigma = 1$$

for all $n \in \mathbb{N}$. It then follows that

$$\lim_{n\to\infty} \frac{\|T_n\|_*}{\operatorname{rank}(T_n)} = \infty.$$

Fortunately, we do not need to rely on (45) for the applications consider in this article. Instead another workaround that uses the notion of coherence, discussed in the next section, is more naturally applicable in our situtations.

VIII. Coherence

We will show in this section that a simple measure of angular constraint called coherence, or rather, the closely related notion of *relative incoherence*, allows us to alleviate two problems simultaneously: the computational intractability of checking for uniqueness discussed in Section IV and the non-existence of a best approximant in Section VI.

Definition 23. Let $\mathbb{H}$ be a Hilbert space provided with scalar product $\langle \cdot, \cdot \rangle$, and let $\Phi \subseteq \mathbb{H}$ be a set of elements of unit norm in $\mathbb{H}$. The **coherence** of Φ is defined as

$$\mu(\Phi) = \sup_{\varphi \neq \psi} |\langle \varphi, \psi \rangle|$$

where the supremum is taken over all distinct pairs $\varphi, \psi \in \Phi$. If $\Phi = \{\varphi_1, \ldots, \varphi_r\}$ is finite, we also write $\mu(\varphi_1, \ldots, \varphi_r) := \max_{p \neq q} |\langle \varphi_p, \varphi_q \rangle|$.

We adopt the convention that whenever we write $\mu(\Phi)$ (resp. $\mu(\varphi_1, \ldots, \varphi_r)$) as in Definition 23, it is implicitly implied that all elements of Φ (resp. $\varphi_1, \ldots, \varphi_r$) are of unit norm.

The notion of coherence has received different names in the literature: mutual incoherence of two dictionaries [27], mutual coherence of two dictionaries [9], the coherence of a subspace projection [8], etc. The version here follows that of [33]. Usually, dictionaries are finite or countable, but we have here a continuum of atoms. Clearly, $0 \leq \mu(\Phi) \leq 1$, and $\mu(\Phi) = 0$ iff $\varphi_1, \ldots, \varphi_r$ are orthonormal. Also, $\mu(\Phi) = 1$ iff Φ contains at least a pair of collinear elements, i.e., $\varphi_p = \lambda \varphi_q$ for some $p \neq q$, $\lambda \neq 0$.

We find it useful to introduce a closely related notion that we call relative incoherence. It allows us to formulate some of our results slightly more elegantly.

Definition 24. Let $\Phi \subseteq \mathbb{H}$ be a set of elements of unit norm. The **relative incoherence** of Φ is defined as

$$\omega(\Phi) = \frac{1 - \mu(\Phi)}{\mu(\Phi)}.$$

For a finite set of unit vectors $\Phi = \{\varphi_1, \ldots, \varphi_r\}$, we will also write $\omega(\varphi_1, \ldots, \varphi_r)$ occasionally.

It follows from our observation about coherence that $0 \leq \omega(\Phi) \leq \infty$, $\omega(\Phi) = \infty$ iff $\varphi_1, \ldots, \varphi_r$ are orthonormal, and $\omega(\Phi) = 0$ iff Φ contains at least a pair of collinear elements.

In the next few subsections, we will see respectively how coherence can inform us about the existence (Section VIII-A), uniqueness (Section VIII-B), as well as both existence and uniqueness (Section VIII-C) of a solution to the best rank-r multilinear approximation problem (28). We will also see how it can be used for establishing exact recoverability (Section VIII-D) and approximation bounds (Section VIII-E) in greedy algorithms.

A. Existence Via Coherence

The goal is to prevent the phenomenon we observed in Example 12 to occur, by imposing natural and weak constraints; we do not want to reduce the search to a compact set. It is clear that the objective is not coercive, which explains why the minimum may not exist. But with an additional condition on the *coherence*, we shall be able to prove existence thanks to coercivity.

The following shows that a solution to the bounded coherence best rank-r approximation problem always exists:

Theorem 25. Let $f \in L^2(X_1 \times \cdots \times X_d)$ be a d-partite function. If

$$\prod_{k=1}^{d} (1 + \omega_k) > r - 1 \tag{46}$$

or equivalently if

$$\prod_{k=1}^{d} \mu_k < \frac{1}{r-1}, \tag{47}$$

where μ_k denotes the coherence as in Definition 23, then

$$\eta = \inf\left\{ \left\| f - \sum_{p=1}^{r} \lambda_p \varphi_{1p} \otimes \cdots \otimes \varphi_{dp} \right\| : \boldsymbol{\lambda} \in \mathbb{C}^r,\ \mu(\varphi_{k1}, \ldots, \varphi_{kr}) \leq \mu_k \right\} \tag{48}$$

$$= \inf\left\{ \left\| f - \sum_{p=1}^{r} \lambda_p \varphi_{1p} \otimes \cdots \otimes \varphi_{dp} \right\| : \boldsymbol{\lambda} \in \mathbb{C}^r,\ \omega(\varphi_{k1}, \ldots, \varphi_{kr}) \geq \omega_k \right\}$$

is attained. Here, $\|\cdot\|$ denotes the L^2-norm on $L^2(X_1 \times \cdots \times X_d)$ and $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_r)$. If desired, we may assume that $\boldsymbol{\lambda} \in \mathbb{R}^r$ and $\lambda_1 \geq \cdots \geq \lambda_r > 0$ by Theorem 2.

Proof: The equivalence between (46) and (47) follows from Definition 24. We show that if either of these conditions are met, then the loss function is coercive. We have the following inequalities

$$\left\| \sum_{p=1}^{r} \lambda_p \varphi_{1p} \otimes \cdots \otimes \varphi_{dp} \right\|^2 = \sum_{p,q=1}^{r} \lambda_p \bar{\lambda}_q \prod_{k=1}^{d} \langle \varphi_{kp}, \varphi_{kq} \rangle$$

$$\geq \sum_{p=1}^{r} \lambda_p \bar{\lambda}_p \prod_{k=1}^{d} \|\varphi_{kp}\|^2 - \sum_{p \neq q}^{r} \left| \lambda_p \bar{\lambda}_q \prod_{k=1}^{d} \langle \varphi_{kp}, \varphi_{kq} \rangle \right|$$

$$\geq \sum_{p=1}^{r} |\lambda_p|^2 - \prod_{k=1}^{d} \mu_k \sum_{p \neq q} |\lambda_p \bar{\lambda}_q| \geq \|\boldsymbol{\lambda}\|_2^2 - (r-1)\|\boldsymbol{\lambda}\|_2^2 \prod_{k=1}^{d} \mu_k$$

where the last inequality follows from

$$\sum_{p \neq q} |\lambda_p \bar{\lambda}_q| \leq (r-1)\|\boldsymbol{\lambda}\|_2^2,$$

which is true because $\sum_{p \neq q} (|\lambda_p| - |\lambda_q|)^2 \geq 0$. This yields

$$\left\| \sum_{p=1}^{r} \lambda_p \varphi_{1p} \otimes \cdots \otimes \varphi_{dp} \right\|^2 \geq \left[1 - (r-1) \prod_{k=1}^{d} \mu_k \right] \|\boldsymbol{\lambda}\|_2^2 \tag{49}$$

Since by assumption $(r-1)\prod_{k=1}^{d} \mu_k < 1$, it is clear that the left hand side of (49) tends to infinity as $\|\boldsymbol{\lambda}\|_2 \to \infty$. And because f is fixed, $\left\| f - \sum_{p=1}^{r} \lambda_p \varphi_{1p} \otimes \cdots \otimes \varphi_{dp} \right\|$ also tends to infinity as $\|\boldsymbol{\lambda}\|_2 \to \infty$. This proves coercivity of the loss function and hence the existential statement. ∎

The condition (46) or, equivalently, (47), in Theorem 25 is sharp in an appropriate sense. Theorem 25 shows that the condition (47) is sufficient in the sense that it guarantees a best rank-r approximation when the condition is met. We show that it is also necessary in the sense that if (47) does not hold, then there are examples where a best rank-r approximation fails to exist.

In fact, let $\hat{f}$ be as in Example VI. As demonstrated in the proof of Theorem VI, the infimum for the case $d = 3$ and $r = 2$,

$$\inf_{\|g_k\| = \|h_k\| = 1,\ \lambda, \mu \geq 0} \|\hat{f} - \lambda g_1 \otimes g_2 \otimes g_3 - \mu h_1 \otimes h_2 \otimes h_3\|$$

is not attained. Since

$$g_k = \frac{\varphi_k + n^{-1}\psi_k}{\|\varphi_k + n^{-1}\psi_k\|}, \quad h_k = \frac{\varphi_k}{\|\varphi_k\|},$$

for $k = 1, 2, 3$, the corresponding coherence

$$\mu(g_k, h_k) \geq |\langle g_k, h_k \rangle| \to 1$$

as $n \to \infty$. For any values of $\mu_1, \mu_2, \mu_3 \in [0, 1]$ such that (47) holds, i.e. $\mu_1\mu_2\mu_3 < 1/(r-1) = 1$, we cannot possibly have $\mu(g_k, h_k) \leq \mu_k$ for all $k = 1, 2, 3$ since

$$\mu(g_1, h_1)\mu(g_2, h_2)\mu(g_3, h_3) \to 1$$

as $n \to \infty$.

B. Uniqueness and Minimality Via Coherence

In order to relate uniqueness and minimality of multilinear decompositions to coherence, we need a simple observation about the notion of Kruskal rank introduced in Definition 7.

Lemma 26. Let $\Phi \subseteq L^2(X_1 \times \cdots \times X_d)$ be finite and $\operatorname{krank} \Phi < \dim \operatorname{span} \Phi$. Then

$$\operatorname{krank} \Phi \geq \frac{1}{\mu(\Phi)}. \tag{50}$$

Proof: Let $s = \operatorname{krank} \Phi + 1$. Then there exists a subset of s distinct unit vectors in Φ, $\{\varphi_1, \ldots, \varphi_s\}$ such that $\alpha_1\varphi_1 + \cdots + \alpha_s\varphi_s = 0$ with $|\alpha_1| = \max\{|\alpha_1|, \ldots, |\alpha_s|\} > 0$. Taking inner product with φ_1 we get $\alpha_1 = -\alpha_2\langle\varphi_2, \varphi_1\rangle - \cdots - \alpha_s\langle\varphi_s, \varphi_1\rangle$ and so $|\alpha_1| \leq (|\alpha_2| + \cdots + |\alpha_s|)\mu(\Phi)$. Dividing by $|\alpha_1|$ then yields $1 \leq (s-1)\mu(\Phi)$. The condition $\operatorname{krank} \Phi < \dim \operatorname{span} \Phi$ prevents Φ from being orthonormal, so $\mu(\Phi) > 0$ and we obtain (50). ∎

We now characterize the uniqueness of the rank-retaining decomposition in terms of coherence introduced in Definition 23.

Theorem 27. Suppose $f \in L^2(X_1 \times \cdots \times X_d)$ has a multilinear decomposition

$$f = \sum_{p=1}^{r} \lambda_p \varphi_{1p} \otimes \cdots \otimes \varphi_{dp}$$

where $\Phi_k := \{\varphi_{k1}, \ldots, \varphi_{kr}\}$ are elements in $L^2(X_k)$ of unit norm and $\operatorname{krank} \Phi_k < \dim \operatorname{span} \Phi_k$ for all $k = 1, \ldots, d$. Let $\omega_k = \omega(\Phi_k)$. If

$$\sum_{k=1}^{d} \omega_k \geq 2r - 1, \tag{51}$$

then $r = \operatorname{rank}(f)$ and the decomposition is essentially unique. In terms of coherence, (51) takes the form

$$\sum_{k=1}^{d} \frac{1}{\mu_k} \geq 2r + d - 1. \tag{52}$$

Proof: Inequality (52) implies that $\sum_{k=1}^{d} \mu_k^{-1} \geq 2r + d - 1$, where μ_k denotes $\mu(\Phi_k)$. If it is satisfied, then so is Kruskal's condition (25) thanks to Lemma 26. The result hence directly follows from Lemma 9 and Definition 24. ∎

Note that unlike the Kruskal ranks in (25), the coherences in (52) are trivial to compute. In addition to uniqueness, an easy but important consequence of Theorem 27 is that it provides

a *readily checkable sufficient condition* for tensor rank, which is NP-hard over any field [42], [43].

Since the purpose of Theorem 27 is to provide a computationally feasible alternative of Lemma 9, excluding the case $\operatorname{krank}\Phi_k = \dim\operatorname{span}\Phi_k$ is not an issue. Note that $\operatorname{krank}\Phi_k = \dim\operatorname{span}\Phi_k$ iff Φ_k comprises linearly independent elements, and the latter can be checked in polynomial time. So this is a case where Lemma 9 can be readily checked and one does not need Theorem 27.

C. Existence and Uniqueness Via Coherence

The following existence and uniqueness sufficient condition may now be deduced from Theorems 25 and 27.

Corollary 28. If $d \geq 3$ and if coherences μ_k satisfy

$$\left(\prod_{k=1}^{d}\mu_k\right)^{1/d} \leq \frac{d}{2r+d-1} \tag{53}$$

then the bounded coherence best rank-r approximation problem has a unique solution up to unimodulus scaling.

Proof: The existence in the case $r = 1$ is assured, because the set of separable functions $\{\varphi_1 \otimes \cdots \otimes \varphi_d : \varphi_k \in L^2(X_k)\}$ is closed. Consider thus the case $r \geq 2$. Since the function $f(x) = \frac{1}{x} - \left(\frac{d}{2x+d-1}\right)^d$ is strictly positive for $x \geq 2$ and $d \geq 3$, condition (53) implies that $\prod_{k=1}^{d}\mu_k$ is smaller than $1/r$, which permits to claim that the solution exists by calling for Theorem 25. Next in order to prove uniqueness, we use the inequality between harmonic and geometric means: if (53) is verified, then we also necessarily have $d\left(\sum_{k=1}^{d}\mu_k^{-1}\right)^{-1} \leq \frac{d}{2r+d-1}$. Hence $\sum_{k=1}^{d}\mu_k^{-1} \geq 2r+d-1$ and we can apply Theorem 27. ∎

In practice, simpler expressions than (53) can be more attractive for computational purposes. These can be derived from the inequalities between means:

$$\left[\frac{1}{d}\sum_{k=1}^{d}\mu_k^{-1}\right]^{-1} \leq \left[\prod_{k=1}^{d}\mu_k\right]^{\frac{1}{d}} \leq \frac{1}{d}\sum_{k=1}^{d}\mu_k \leq \left[\frac{1}{d}\sum_{k=1}^{d}\mu_k^2\right]^{\frac{1}{2}}.$$

Examples of stronger sufficient conditions that could be used in place of (53) include

$$\sum_{k=1}^{d}\mu_k \leq \frac{d^2}{2r+d-1}, \tag{54}$$

$$\sum_{k=1}^{d}\mu_k^2 \leq d\left(\frac{d}{2r+d-1}\right)^2. \tag{55}$$

Another simplification can be performed, which yields differentiable expressions of the constraints if (55) is to be used. In fact, noting that for any set of numbers $x_1, \ldots, x_n \in \mathbb{C}$, $\max_{i=1,\ldots,n}|x_i| \leq \sqrt{\sum_{i=1}^{n}|x_i|^2}$, a sufficient condition ensuring that (55) is satisfied, and hence (53), is

$$\sum_{k=1}^{d}\sum_{p<q}|\langle\varphi_{kp},\varphi_{kq}\rangle|^2 \leq d\left(\frac{d}{2r+d-1}\right)^2.$$

D. Exact Recoverability Via Coherence

We now describe a result that follows from the remarkable work of Temlyakov. It allows us to in principle determine the multilinear decomposition meeting the type of coherence conditions in Section VIII-A.

Some additional notations would be useful. We let $\Phi \subseteq \{f \in L^2(X_1 \times \cdots \times X_d : \operatorname{rank}(f) = 1\}$ be a *dictionary*[6] of separable functions (i.e. rank-1) in $L^2(X_1 \times \cdots \times X_d)$ that meets a bounded coherence condition, i.e.

$$\mu(\Phi) < \mu \tag{56}$$

for some $\mu \in [0, 1)$ to be chosen later. Recall that the elements of Φ are implicitly assumed to be of unit norm (cf. remark after Definition 23).

Let $t \in (0, 1]$. The *weakly orthogonal greedy algorithm* (WOGA) is simple to describe: Set $f_0 = f$. For each $m \in \mathbb{N}$, we inductively define a sequence of f_m's as follows:

1) $g_m \in \Phi$ is any element satisfying
$$|\langle f_{m-1}, g_m\rangle| \geq t \sup_{g\in\Phi}|\langle f_{m-1}, g\rangle|;$$
2) $h_m \in L^2(X_1 \times \cdots \times X_d)$ is a projection of f onto $\operatorname{span}(g_1, \ldots, g_m)$, i.e.
$$h_m \in \operatorname{argmin}\{\|f - g\| : g \in \operatorname{span}(g_1, \ldots, g_m)\}; \tag{57}$$
3) $f_m \in L^2(X_1 \times \cdots \times X_d)$ is a deflation of f by h_m, i.e.
$$f_m = f - h_m.$$

Note that deflation alone, without the coherence requirement, generally does not work for computing multilinear decompositions [65]. The following result, adapted here for our purpose, was proved for any arbitrary dictionary in [69].

Theorem 29 (Temlyakov). Suppose $f \in L^2(X_1 \times \cdots \times X_d)$ has a multilinear decomposition

$$f = \sum_{p=1}^{r}\lambda_p\varphi_{1p} \otimes \cdots \otimes \varphi_{dp}$$

with $\varphi_{1p} \otimes \cdots \otimes \varphi_{dp} \in \Phi$ and the condition that

$$r < \frac{t}{1+t}\left(1 + \frac{1}{\mu}\right)$$

for some $t \in (0, 1]$. Then the WOGA algorithm recovers the factors exactly, or more precisely, $f_r = 0$ and thus $f = h_r$.

So h_r, by its definition in (57) and our choice of Φ, is given in the form of a linear combination of rank-1 functions, i.e., an rank-r multilinear decomposition.

E. Greedy Approximation Bounds Via Coherence

This discussion in Section VIII-D pertains to exact recovery of a rank-r multilinear decomposition although our main problem really takes the form of a best rank-r approximation more often than not. We will describe some greedy approximation bounds for the approximation problem in this section.

[6] A *dictionary* is any set $\Phi \subseteq \mathbb{H}$ whose linear span is dense in the Hilbert space $\mathbb{H}$.

We let

$$\sigma_r(\hat{f}) := \inf_{\boldsymbol{\alpha}\in\mathbb{C}^r,\ \|\varphi_{kp}\|=1} \left\| \hat{f} - \sum_{p=1}^{r} \alpha_p \prod_{k=1}^{d} \varphi_{kp} \right\|.$$

By our definition of rank and border rank,

$$\begin{aligned}\sigma_r(\hat{f}) &= \inf\{\|\hat{f} - f\| : \operatorname{rank}(f) \le r\} \\ &= \min\{\|\hat{f} - f\| : \overline{\operatorname{rank}}(f) \le r\}.\end{aligned}$$

It would be wonderful if greedy algorithms along the lines of what we discussed in Section VIII-D could yield an approximant within some provable bounds that is a factor of $\sigma_r(\hat{f})$. However this is too much to hope for mainly because a dictionary comprising all separable functions, i.e., $\{f : \operatorname{rank}(f) = 1\}$ is far too large to be amenable to such analysis. This does not prevent us from considering somewhat more restrictive dictionaries like what we did in the previous section. So again, let $\Phi \subseteq \{f \in L^2(X_1 \times \cdots \times X_d) : \operatorname{rank}(f) = 1\}$ be such that

$$\mu(\Phi) < \mu$$

for some given $\mu \in [0, 1)$ to be chosen later. Let us instead define

$$s_r(\hat{f}) = \inf_{\boldsymbol{\alpha}\in\mathbb{C}^r,\ \varphi_p\in\Phi} \left\| \hat{f} - \sum_{p=1}^{r} \alpha_p \varphi_p \right\|.$$

Clearly

$$\sigma_r(\hat{f}) \le s_r(\hat{f}) \tag{58}$$

since the infimum is taken over a smaller dictionary.

The special case where $t = 1$ in the WOGA described in Section VIII-D is also called the *orthogonal greedy algorithm* (OGA). The result we state next comes from the work of a number of people done over the last decade: (59) is due to Gilbert, Muthukrisnan, and Strauss in 2003 [32]; (60) is due to Tropp in 2004 [72]; (61) is due to Dohono, Elad, and Temlyakov in 2006 [28]; and (62) is due to Livshitz in 2012 [50]. We merely apply these results to our approximation problem here.

Theorem 30. Let $\hat{f} \in L^2(X_1 \times \cdots \times X_d)$ and f_r be the rth iterate as defined in WOGA with $t = 1$ and input $\hat{f}$.

(i) If $r < \frac{1}{32}\mu^{-1}$, then

$$\|\hat{f} - f_r\| \le 8r^{1/2} s_r(\hat{f}). \tag{59}$$

(ii) If $r < \frac{1}{3}\mu^{-1}$, then

$$\|\hat{f} - f_r\| \le (1 + 6r)^{1/2} s_r(\hat{f}). \tag{60}$$

(iii) If $r \le \frac{1}{20}\mu^{-2/3}$, then

$$\|\hat{f} - f_{r\log r}\| \le 24 s_r(\hat{f}). \tag{61}$$

(iv) If $r \le \frac{1}{20}\mu^{-1}$, then

$$\|\hat{f} - f_{2r}\| \le 3 s_r(\hat{f}). \tag{62}$$

It would be marvelous if one could instead establish bounds in (59), (60), (61), and (62) with $\sigma_r(\hat{f})$ in place of $s_r(\hat{f})$ and $\{f : \operatorname{rank}(f) = 1\}$ in place of Φ, dropping the coherence μ altogether. In which case one may estimate how well the rth OGA iterates f_r approximates the best rank-r approximation. This appears to be beyond present capabilites.

We would to note that although the approximation theoretic technqiues (coherence, greedy approximation, redundant dictionaries, etc) used in this article owe their newfound popularity to compressive sensing, they owe their roots to works of the Russian school of approximation theorists (e.g., Boris Kashin, Vladimir Temlyakov, et al.) dating back to the 1980s. We refer readers to the bibliography of [68] for more information.

F. Checking Coherence-Based Conditions

Since the conditions in Theorems 25, 27, 29, and Corollary 28 all involve coherence, we will say a brief word about its computation.

It has recently been established that computing the spark of a finite set of vectors Φ, i.e., the size of the smallest linearly dependent subset of Φ, is strongly NP-hard [70]. Since $\operatorname{spark}\Phi = \operatorname{krank}\Phi + 1$, it immediately follows that the same is true for Kruskal rank.

Corollary 31 (Kruskal rank is NP-hard) Let $\mathbb{H}$ be a Hilbert space and $\Phi \subseteq \mathbb{H}$ be finite. The optimization problem of computing

$$\operatorname{krank}(\Phi) := \max\left\{k : \text{all } \Psi \in \binom{\Phi}{k} \text{ linearly independent}\right\}$$

is strongly NP-hard. $\binom{\Phi}{k}$ = set of all k-element subsets of Φ.

Given the NP-hardness of Kruskal rank, one expects that Lemma 9, as well as its finite-dimensional counterparts [41], [61], would be computationally intractable to apply in reality. The reciprocal of coherence is therefore a useful surrogate for Kruskal rank by virtue of Lemma 26 and the fact that computing $\mu(\Phi)$ requires only $r(r-1)/2$ inner products $\langle\varphi_i, \varphi_j\rangle$, $i \ne j$, where $r = |\Phi|$.

For finite-dimensional problems, i.e., $X_1, \ldots, X_d$ are all finite sets, we may regard Φ as a matrix in $\mathbb{C}^{n\times r}$, and $\mu(\Phi)$ is simply the largest entry in magnitude in the Gram matrix $\Phi^{\mathsf{T}}\Phi$, which may be rapidly computed using Strassen-type algorithms [67] in numerically stable ways [23]. For infinite-dimensional problems, the problem depends on the cost of integrating complex-valued functions defined on X_k, $k = 1, \ldots, d$. For example, if $X_1, \ldots, X_d$ are all finite intervals of $\mathbb{R}$, one may use *inner product quadratures* [13] to efficiently compute the Gram matrix $(\langle\varphi_i, \varphi_j\rangle)_{i,j=1}^n \in \mathbb{R}^{n\times n}$ and thereby find $\mu(\Phi)$.

IX. Applications

Various applications, many under the headings[7] of CANDECOMP [11] and PARAFAC [36], have appeared in psychometrics and, more recently, also other data analytic applications. We found that many of these applications suffer from a regretable defect — there are no compelling reasons nor rigorous arguments that support the use of a rank-r multilinear decomposition model. The mere fact that a data set may be

[7]Other than CANDECOMP and PARAFAC, the finite-dimensional multilinear decompositions have also been studied under the names CP, CAND, canonical decomposition, and canonical polyadic decompositions.

cast in the form of a d-dimensional array $A \in \mathbb{C}^{n_1 \times \cdots \times n_d}$ does not mean that (13) would be the right or even a reasonable thing to do. In particular, how would one interpret the factors φ_{kp}'s when $d > 2$? When $d = 2$, one could arguably interpret these as principal or varimax components when orthonormality is imposed but for general $d > 2$, a convincing application of a model based on the rank-r multilinear decomposition (13) must rely on careful arguments that follow from first principles.

The goal of this section is two-fold. First we provide a selection of applications where the rank-r multilinear decomposition (13) arises naturally via considerations of first principles (in electrodynamics, quantum mechanics, wave propagation, etc). Secondly, we demonstrate that the coherence conditions discussed extensively in Section VIII invariably have reasonable interpretations in terms of physical quantities.

The use of a rank-r multilinear decomposition model in signal processing via higher-order statistics has a long history [10], [16], [31], [59], [60]. Our signal processing applications here are of a different nature, they are based on geometrical properties of sensor arrays instead of considerations of higher-order statistics. This line of argument first appeared in the work of Sidiropoulos and Bro [62], which is innovative and well-motivated by first principles. However, like all other applications considered thus far, whether in data analysis, signal processing, psychometrics, or chemometrics, it does not address the serious nonexistence problem that we discussed at length in Section VIII-A. Without any guarantee that a solution to (28) exists, one can never be sure when the model would yield a solution. Another issue of concern is that the Kruskal uniqueness condition in Lemma 9 has often been invoked to provide evidence of a unique solution but we now know that this condition is practically impossible to check because of Corollary 31. The applications considered below would use the coherence conditions developed in Section VIII to avoid these difficulties. More precisely, Theorem 25, Theorem 27, and Corollary 28 are invoked to guarantee the existence of a solution to the approximation problem and provide readily checkable conditions for uniqueness of the solution, all via the notion of coherence. Note that unlike Kruskal's condition, which applies only to an *exact* decomposition, Corollary 28 gives uniqueness of an approximation in noisy circumstances.

In this section, applications are presented in finite dimension. In order to avoid any confusion, X^*, X^{H} and X^{T} will denote complex conjugate, hermitian transpose, and transpose, of the matrix X respectively.

A. Joint Channel and Source Estimation

Consider a narrow band transmission problem in the far field. We assume here that we are in the context of wireless telecommunications, but the same principle could also apply in other areas. Let r signals impinge on an array, so that their mixture is recorded. We wish to recover the original signals and to estimate their directions of arrival and respective powers at the receiver. If the channel is specular, some of these signals can correspond to different propagation paths of the same radiating source, and are therefore correlated. In other words, r does not denote the number of sources, but the total number of distinct paths viewed from the receiver.

In the present framework, we assume that channels can be time-varying, but that they can be regarded to be constant over a sufficiently short observation length. The goal is to be able to work with extremely short samples.

In order to face this challenge, we assume that the sensor array is structured, as in [62]. More precisely, the sensor array comprises a *reference array* with n_1 sensors, whose location is defined by a vector $\mathbf{b}_i \in \mathbb{R}^3$, and $n_2 - 1$ other subarrays obtained from the reference array by a translation in space defined by a vector $\mathbf{\Delta}_j \in \mathbb{R}^3$, $j = 2, \ldots, n_2$. The reference subarray is numbered with $j = 1$ in the remainder.

Under these assumptions, the signal received at discrete time t_k, $k = 1, \ldots, n_3$, on the ith sensor of the reference subarray can be written as

$$s_{i,1}(k) = \sum_{p=1}^{r} \sigma_p(t_k) \exp(\psi_{i,p})$$

with $\psi_{i,p} = \jmath \frac{\omega}{C}(\mathbf{b}_i^{\mathsf{T}} \mathbf{d}_p)$ where the dotless $\jmath$ denotes $\sqrt{-1}$; $\mathbf{d}_p \in \mathbb{R}^3$ is of unit norm and denotes direction of arrival of the pth path, C denotes the wave celerity, and ω denotes the pulsation. Next, on the jth subarray, $j = 2, \ldots, n_2$, we have

$$s_{i,j}(k) = \sum_{p=1}^{r} \sigma_p(t_k) \exp(\psi_{i,j,p}) \tag{63}$$

with $\psi_{i,j,p} = \jmath \frac{\omega}{C}(\mathbf{b}_i^{\mathsf{T}} \mathbf{d}_p + \mathbf{\Delta}_j^{\mathsf{T}} \mathbf{d}_p)$. If we let $\mathbf{\Delta}_1 = \mathbf{0}$, then (63) also applies to the reference subarray. The crucial feature of this structure is that variables i and j decouple in the function $\exp(\psi_{i,j,p})$, yielding a relation resembling the rank-retaining multilinear decomposition:

$$s_{i,j}(k) = \sum_{p=1}^{r} \lambda_p u_{ip} v_{jp} w_{kp}$$

where $u_{ip} = \exp\left(\jmath \frac{\omega}{C} \mathbf{b}_i^{\top} \mathbf{d}_p\right)$, $v_{jp} = \exp\left(\jmath \frac{\omega}{C} \mathbf{\Delta}_j^{\mathsf{T}} \mathbf{d}_p\right)$ and $w_{kp} = \sigma_p(t_k)/\|\boldsymbol{\sigma}_p\|$, $\lambda_p = \|\boldsymbol{\sigma}_p\|$.

By computing a rank-retaining decomposition of the hypermatrix $S = (s_{i,j}(k)) \in \mathbb{C}^{n_1 \times n_2 \times n_3}$, one may jointly estimate: (i) signal waveforms $\sigma_p(k)$, and (ii) directions of arrival $\mathbf{d}_p$ of each propagation path, provided $\mathbf{b}_i$ or $\mathbf{\Delta}_j$ are known.

However, the observation model (63) is not realistic, and an additional error term should be added in order to account for modeling inaccuracies and background noise. It is customary (and realistic thanks to the central limit theorem) to assume that this additive error has a continuous probability distribution, and that therefore the hypermatrix S has the *generic rank*. Since the generic rank is at least as large as $\lceil n_1 n_2 n_3/(n_1 + n_2 + n_3 - 2) \rceil$, which is always larger than Kruskal's bound [19], we are led to the problem of approximating the hypermatrix S by another of rank r. We have seen that the angular constraint imposed in Section VIII permits us to deal with a well-posed problem. In order to see the physical meaning of this constraint, we need to first define the tensor product between sensor subarrays.

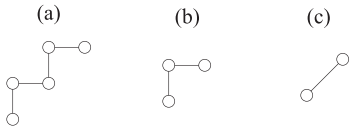


Fig. 1. Antenna array (a) obtained as tensor product of subarrays (b) & (c).

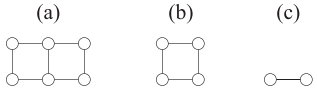


Fig. 2. Antenna array (a) obtained as tensor product of subarrays (b) & (c).

B. Tensor Product Between Sensor Subarrays

The sensor arrays we encounter are structured, in the sense that the whole array is generated by one subarray defined by the collection of vector locations $\{\mathbf{b}_i \in \mathbb{R}^3 : 1 \leq i \leq n_1\}$, and a collection of translations in space, $\{\mathbf{\Delta}_j \in \mathbb{R}^3 : 1 \leq j \leq n_2\}$. If we define vectors

$$\begin{aligned} \mathbf{u}_p &= \frac{1}{\sqrt{n_1}} \left[\exp\left(\jmath \frac{\omega}{C} \mathbf{b}_i^{\mathsf{T}} \mathbf{d}_p\right)\right]_{i=1}^{n_1}, \\ \mathbf{v}_p &= \frac{1}{\sqrt{n_2}} \left[\exp\left(\jmath \frac{\omega}{C} \mathbf{\Delta}_j^{\mathsf{T}} \mathbf{d}_p\right)\right]_{j=1}^{n_2}, \\ \mathbf{w}_p &= \boldsymbol{\sigma}_p / \|\boldsymbol{\sigma}_p\|, \end{aligned} \tag{64}$$

then we may view all measurements as the superimposition of decomposable hypermatrices $\lambda_p \mathbf{u}_p \otimes \mathbf{v}_p \otimes \mathbf{w}_p$.

Geometrical information of the sensor array is contained in $\mathbf{u}_p \otimes \mathbf{v}_p$ while energy and time information on each path p is contained in λ_p and $\mathbf{w}_p$ respectively. Note that the reference subarray and the set of translations play symmetric roles, in the sense that $\mathbf{u}_p$ and $\mathbf{v}_p$ could be interchanged without changing the whole array. This will become clear with a few examples.

When we are given a structured sensor array, there can be several ways of splitting it into a tensor product of two (or more) subarrays, as shown in the following simple examples.

Example 32. Define the matrix of sensor locations

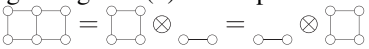
$$[\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3] = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

This subarray is depicted in Figure 1(b). By translating it via the translation in Figure 1(c) one obtains another subarray. The union of the two subarrays yields the array of Figure 1(a). The same array is obtained by interchanging roles of the two subarrays, i.e., three subarrays of two sensors deduced from each other by two translations.

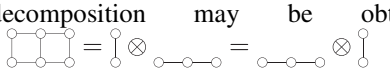
Example 33. Define the array by

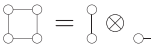
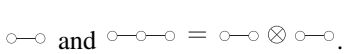
$$[\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_6] = \begin{bmatrix} 0 & 1 & 2 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}. \tag{65}$$

This array, depicted in Figure 2(a), can either be obtained from the union of subarray of Figure 2(b) and its translation defined by Figure 2(c), or from the array of Figure 2(c) translated three times according to Figure 2(b). We express this relationship as



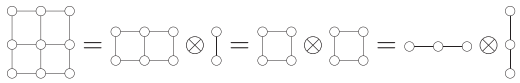
Another decomposition may be obtained as



In fact,  and  . However, it is important to stress that the various decompositions of the whole array into tensor products of subarrays are not equivalent from the point of view of performance. In particular, the Kruskal bound can be different, as we will see next.

Similar observations can be made for grid arrays in general.

Example 34. Take an array of 9 sensors located at $(x, y) \in \{1, 2, 3\} \times \{1, 2, 3\}$. We have the relations

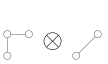
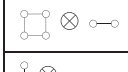
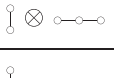
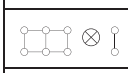
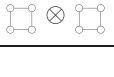


among others.

Let us now have a look at the maximal number of sources $r_{\max}$ that can be extracted from a $n_1 \times n_2 \times n_3$ hypermatrix in the absence of noise. A sufficient condition is that the total number of paths, r, is smaller than Kruskal's bound (25). We shall simplify the bound by making two assumptions: (a) the loading matrices are generic, i.e., they are of full rank, and (b) the number of paths is larger than the sizes n_1 and n_2 of the two subarrays entering the array tensor product, and smaller than the number of time samples, n_3. Under these simplifying assumptions, Kruskal's bound becomes $2r_{\max} \leq n_1 + n_2 + r_{\max} - 2$, or:

$$r_{\max} = n_1 + n_2 - 2 \tag{66}$$

The table below illustrates the fact that the choice of subarrays has an impact on this bound.

Array	Subarray product	n_1	n_2	$r_{\max}$
		3	2	3
		4	2	4
		2	3	3
		3	3	4
		6	2	6
		4	4	6

C. Significance of the Angular Constraint

We are now in a position to interpret the meanings of the various coherences in light of this application. According to the notations given in (64), the first coherence

$$\mu_1 = \max_{p \neq q} |\mathbf{u}_p^{\mathsf{H}} \mathbf{u}_q|$$

corresponds to the angular separation viewed from the reference subarray. The vectors $\mathbf{b}_i$ and $\mathbf{d}_p$ have unit norm, as do the vectors $\mathbf{u}_p$. The quantity $|\mathbf{u}_p^{\mathsf{H}} \mathbf{u}_q|$ may thus be viewed as a

measure of angular separation between $\mathbf{d}_p$ and $\mathbf{d}_q$, as we shall demonstrate in Proposition IX-C.

Definition 35. We shall say that a collection of vectors $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ is **resolvent** with respect to a direction $\mathbf{v} \in \mathbb{R}^3$ if there exist two indices k and l such that $\mathbf{v} = \mathbf{b}_k - \mathbf{b}_l$ and

$$0 < \|\mathbf{v}\| < \frac{\lambda}{2},$$

where $\lambda = 2\pi C/\omega$ denotes the wavelength.

Let $\mathbf{b}_i$, $\mathbf{d}_p$ and $\mathbf{u}_q$ be defined as in (64), $i = 1, \ldots, n_1$, $p, q = 1, \ldots, n_2$. Then we have the following.

Proposition 36. If $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ is resolvent with respect to three linearly independent directions, then

$$|\mathbf{u}_p^{\mathsf{H}} \mathbf{u}_q| = 1 \quad \Leftrightarrow \quad \mathbf{d}_p = \mathbf{d}_q.$$

Proof: Assume that $|\mathbf{u}_p^{\mathsf{H}} \mathbf{u}_q| = 1$. Since they are of unit norm, vectors $\mathbf{u}_p$ and $\mathbf{u}_q$ are collinear with a unit modulus proportionality factor. Hence, from (64), for all $j, k = 1, \ldots, n_1$, $(\mathbf{b}_j - \mathbf{b}_k)^{\mathsf{T}}(\mathbf{d}_p - \mathbf{d}_q) \in \lambda\mathbb{Z}$, where λ is as in Definition 25. Since $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ is resolvent, there exist k, l such that $0 < \|\mathbf{b}_k - \mathbf{b}_l\| < \lambda/2$. As the vectors $\mathbf{d}_p$ are of unit norm, $\|\mathbf{d}_p - \mathbf{d}_q\| \leq 2$ and we necessarily have that $(\mathbf{b}_k - \mathbf{b}_l)^{\mathsf{T}}(\mathbf{d}_p - \mathbf{d}_q) = 0$, i.e., $\mathbf{d}_p - \mathbf{d}_q$ is orthogonal to $\mathbf{b}_k - \mathbf{b}_l$. The same reasoning can be carried out with the other two independent vectors. The vector $\mathbf{d}_p - \mathbf{d}_q$ must be $\mathbf{0}$ because it is orthogonal to three linearly independent vectors in $\mathbb{R}^3$. The converse is immediate from the definition of $\mathbf{u}_q$. ∎

Note that the condition in Definition 35 is not very restrictive, since sensor arrays usually contain sensors separated by half a wavelength or less. Thanks to Proposition 36, we now know that uniqueness of the matrix factor $U = [\mathbf{u}_1, \ldots, \mathbf{u}_r]$ and the identifiability of the directions of arrival $\mathbf{d}_p$ are equivalent. By the results of Section VIII, the uniqueness can be ensured by a constraint on coherence such as (53).

As in Section IX-B, the second coherence may be interpreted as a measure of the minimal angular separation between paths, viewed from the subarray defining translations.

The third coherence is the maximal correlation coefficient between signals received from various paths on the array

$$\mu_3 = \max_{p \neq q} \frac{|\boldsymbol{\sigma}_p^{\mathsf{H}} \boldsymbol{\sigma}_q|}{\|\boldsymbol{\sigma}_p\| \|\boldsymbol{\sigma}_q\|}.$$

In conclusion, the best rank-r approximation exists and is unique if either signals propagating through various paths are not too correlated, or if their direction of arrival are not too close, where "not too" is taken to mean that the product of coherences satisfies inequality (53) of Corollary 28. In other words, one can separate paths with arbitrarily high correlation provided they are sufficiently well separated in space.

Hence, the decomposition of a sensor array into a tensor product of two (or more) sensor subarrays depends not only on Kruskal's bound, as elaborated in Section IX-B, but also on the ability of the latter subarrays to separate two distinct directions of arrival (cf. Proposition 36).

D. CDMA Communications

The application to antenna array processing we described in Section IX-A also applies to all source separation problems [18], provided an additional diversity is available. An example is the case of Code Division Multiple Access (CDMA) communications. In fact, as pointed out in [63], it is possible to distinguish between symbol and chip diversities. We will elaborate on the latter example.

Consider a downlink CDMA communication with r users, each assigned a spreading sequence $C_p(k)$, $p = 1, \ldots, r$, $k = 1, \ldots, n$. Denote by A_{ip} the complex gain between sensor i, $i = 1, \ldots, m$, and user p, by S_{jp} the symbol sequence transmitted by user p, $j \in \mathbb{Z}$, and by $H_p(k)$ the channel impulse response of user p. The signal received on sensor i during the kth chip of the jth symbol period takes the form

$$T_{ijk} = \sum_{p=1}^{r} A_{ip} S_{jp} B_{kp}$$

where $B_{kp} = \sum_t H_p(k-t) C_p(t)$ denotes the output of the pth channel excited by the pth coding sequence, upon removal of the guard chips (which may be affected by two different consecutive symbols) [63].

The columns of matrix $B = (B_{kp}) \in \mathbb{R}^{n \times r}$ are often referred to as "effective codes", and coincide with spreading codes if the channel is memoryless and noiseless. In practice, the receiver filter is matched to the transmitter shaping filter combined with the propagation channel, so that effective and spreading codes are ideally proportional. Under these conditions, the coherence μ_C accounts for the angular separation between spreading sequences: $\mu_C = 0$ means that they are orthogonal. On the other hand, $\mu_A = 0$ means that the symbol sequences are all uncorrelated. Lastly, as seen in Proposition 36, $\mu_B = 1$ means that the directions of arrival are collinear.

In order to avoid multiple access interferences, spreading sequences are usually chosen to be uncorrelated for all delays, which implies that they are orthogonal. However, the results obtained in Section VIII show that spreading sequences *do not need to be orthogonal*, and symbol sequences *need not be uncorrelated*, as long as the directions of arrival are not collinear. In particular, shorter spreading sequences may be used for the same number of users, which increases throughput. Alternatively, for a given spreading gain, one may increase the number of users. These are possible because the coherence conditions in Section VIII allow one to relax the constraint of having almost orthogonal spreading sequences. On the other hand, some directions of arrival may be collinear if the corresponding spreading sequences are sufficiently well separated angularly. These conclusions are essentially valid when users are synchronized, i.e., for downlink communications.

E. Polarization

The use of polarization as an additional diversity has its roots in [56]. Several attempts to use this diversity in the framework of tensor-based source localization and estimation can be found in the literature [35].

In this framework, we consider again an array of n_1 sensors, whose locations are given by $\mathbf{b}_i \in \mathbb{R}^3$, $i = 1, \ldots, n$. We assume a narrow-band transmission in the far field (i.e., sources, or source paths, are all seen as plane waves at the receiver sensor array). The difference with Section IX-B is that translation diversity is not mandatory anymore, provided that the impinging waves are polarized and that their polarization is neither linear nor circular. One measures the electric and magnetic fields at each sensor as a function of time, so that $n_2 = 6$. More precisely, $\mathbf{v}_p$ of (64) is replaced by

$$\mathbf{v}_p = B_p \mathbf{g}_p \tag{67}$$

where $B_p \in \mathbb{R}^{6\times 2}$ depends only on the direction of arrival $\mathbf{d}_p$ (defined in Section IX-B), and $\mathbf{g}_p \in \mathbb{C}^2$ depends only on the orientation and ellipticity of the polarization of the pth wave.

Coherences μ_1 and μ_3 are the same as in Section IX-C, and represent respectively the angular separation between directions of arrival, and correlation between arriving sources. It is slightly more difficult to see the significance of μ_2, the coherence associated with polarization.

For this, we need to go into more details [56]. Let $\alpha_p \in (-\pi/2, \pi/2]$ and $\beta_p \in (-\pi/4, 0) \cup (0, \pi/4)$ denote respectively the orientation and ellipticity angles of the polarization of the pth wave. Let $\theta_p \in [0, 2\pi)$ and $\phi_p \in (-\pi/2, \pi/2]$ denote respectively the azimuth and elevation of the direction of arrival of the pth path. We have

$$B_p = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{e}_p & \mathbf{f}_p \\ \mathbf{f}_p & -\mathbf{e}_p \end{bmatrix}, \qquad \mathbf{g}_p = Q(\alpha_p)\mathbf{h}_p,$$

where

$$\mathbf{e}_p = \begin{bmatrix} -\sin\theta_p \\ \cos\theta_p \\ 0 \end{bmatrix}, \qquad \mathbf{f}_p = \begin{bmatrix} -\cos\theta_p \sin\phi_p \\ -\sin\theta_p \sin\phi_p \\ \cos\phi_p \end{bmatrix},$$

$$Q(\alpha) = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix}, \qquad \mathbf{h}_p = \begin{bmatrix} \cos\beta_p \\ \jmath \sin\beta_p \end{bmatrix}.$$

The unit vector defining the pth direction of arrival is

$$\mathbf{d}_p = \begin{bmatrix} \cos\theta_p \cos\phi_p \\ \sin\theta_p \cos\phi_p \\ \sin\phi_p \end{bmatrix}.$$

So the triplet $(\mathbf{d}_p, \mathbf{e}_p, \mathbf{f}_p)$ forms a right orthonormal triad.

Lemma 37. $|\mathbf{g}_p^{\mathsf{H}} \mathbf{g}_q| = 1$ if and only if $\alpha_p = \alpha_q + k\pi$ and $\beta_p = \beta_q$, $k \in \mathbb{Z}$.

Proof: First note that $Q(\alpha_p)^{\mathsf{H}} Q(\alpha_q) = Q(\alpha_q - \alpha_p)$. Hence $\mathbf{g}_p^{\mathsf{H}} \mathbf{g}_q$ can be of unit modulus only if $\mathbf{h}_p$ and $Q(\alpha_q - \alpha_p)\mathbf{h}_q$ are collinear. But the first entry of $\mathbf{h}_p$ is real and the second is purely imaginary. So the corresponding imaginary and real parts of $Q(\alpha_q - \alpha_p)\mathbf{h}_q$ must be zero, which implies that $\sin(\alpha_q - \alpha_p) = 0$. Consequently $Q(\alpha_q - \alpha_p) = \pm I$, which yields $\mathbf{h}_p = \pm \mathbf{h}_q$. But because the angle β lies in the interval $(-\pi/4, \pi/4)$, only the positive sign is acceptable. ∎

Proposition 38. $|\mathbf{v}_p^{\mathsf{H}} \mathbf{v}_q| \leq 1$, with equality if and only if $\alpha_p = \alpha_q + k\pi$, $\beta_p = \beta_q$, $\theta_p = \theta_q + k'\pi$ and $\phi_p = \phi_q$, $k, k' \in \mathbb{Z}$. *Proof:* We have $|\mathbf{v}_p^{\mathsf{H}} \mathbf{v}_q| = |\mathbf{g}_p^{\mathsf{H}} B_p^{\mathsf{T}} B_q \mathbf{g}_q|$. Notice that the matrix $B_p^{\mathsf{T}} B_q$ is of the form

$$B_p^{\mathsf{T}} B_q = \begin{bmatrix} \gamma & \eta \\ -\eta & \gamma \end{bmatrix}$$

where γ and η are real, $\gamma = \frac{1}{2}(\mathbf{e}_p^{\mathsf{T}} \mathbf{e}_q + \mathbf{f}_p^{\mathsf{T}} \mathbf{f}_q)$ and $\eta = \frac{1}{2}(\mathbf{e}_p^{\mathsf{T}} \mathbf{f}_q - \mathbf{f}_p^{\mathsf{T}} \mathbf{e}_q)$. Since $\mathbf{g}_p$ and $\mathbf{g}_q$ are of unit norm, $|\mathbf{v}_p^{\mathsf{H}} \mathbf{v}_q|$ can be of unit modulus only if $B_p^{\mathsf{T}} B_q$ has an eigenvalue of unit modulus, which requires that $\gamma^2 + \eta^2 = 1$. We now prove that $\gamma^2 + \eta^2 \leq 1$ with equality if and only if the four sets of equalities hold.

With this goal in mind, define the 6-dimensional vectors

$$\mathbf{z} = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{e}_p \\ \mathbf{f}_p \end{bmatrix}, \quad \mathbf{w} = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{e}_q \\ \mathbf{f}_q \end{bmatrix}, \quad \mathbf{w}' = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{f}_q \\ -\mathbf{e}_q \end{bmatrix}.$$

Then $\gamma = \mathbf{z}^{\mathsf{T}} \mathbf{w}$ and $\gamma = \mathbf{z}^{\mathsf{T}} \mathbf{w}'$. Decompose $\mathbf{z}$ into two orthogonal parts: $\mathbf{z} = \mathbf{z}_0 + \mathbf{z}_1$, with $\mathbf{z}_0 \in \text{span}\{\mathbf{w}, \mathbf{w}'\}$ and $\mathbf{z}_0 \perp \mathbf{z}_1$. Clearly, $\gamma^2 + \eta^2 = \|\mathbf{z}_0\|^2$. Moreover, $\|\mathbf{z}_0\|^2 \leq \|\mathbf{z}\|^2 = 1$, with equality if and only if $\mathbf{z} \in \text{span}\{\mathbf{w}, \mathbf{w}'\}$. By inspection of the definitions of $\mathbf{e}_p$ and $\mathbf{e}_q$, we see that the third entry of $\mathbf{z}$ and $\mathbf{w}$ is $\mathbf{0}$. Hence $\mathbf{z} \in \text{span}\{\mathbf{w}, \mathbf{w}'\}$ is possible only if either $\mathbf{z}$ is collinear to $\mathbf{w}$ or if the third entry of $\mathbf{w}'$ is $\mathbf{0}$. In the latter case, it means that $\phi_q = \pi/2$, and so $\phi_p = \pi/2$ and $\theta_p = \theta_q$. In the former case, it can be seen that $\sin\theta_p = \sin\theta_q$, and finally that $\phi_p = \phi_q$.

The last step is to rewrite γ and η as a function of angle $\theta_p - \theta_q$, using trigonometric relations: $\gamma = \cos(\theta_p - \theta_q)(1 + \sin\phi_p \sin\phi_q) + \cos\phi_p \cos\phi_q$ and $\eta = \sin(\theta_p - \theta_q)(\sin\phi_p + \sin\phi_q)$. This eventually shows that $\gamma = 1$ and $\eta = 0$. As a consequence, $|\mathbf{v}_p^{\mathsf{H}} \mathbf{v}_q| = 1$ only if $B_p^{\mathsf{T}} B_q = I$, and the result follows from Lemma 37. ∎

Proposition 38 shows that a constraint on the coherence μ_2 compels source paths to have either different directions of arrival or different polarizations, giving μ_2 physical meaning.

F. Fluorescence Spectral Analysis

Here is a well-known application to fluorescence spectral analysis originally discussed in [64]. We use the notations in Example 6. Suppose we have l samples with an unknown number of pure substances in different concentrations that are fluorescent. If a_{ijk} is the measured fluorescence emission intensity at wavelength λ_j^{em} of the ith sample excited with light of wavelength λ_k^{ex}. The measured data is a 3-dimensional hypermatrix $A = (a_{ijk}) \in \mathbb{R}^{l\times m\times n}$. At low concentrations, Beer's law of spectroscopy (which is in turn a consequence of fundamental principles in quantum mechanics) can be linearized [51], yielding a rank-retaining decomposition

$$A = \mathbf{x}_1 \otimes \mathbf{y}_1 \otimes \mathbf{z}_1 + \cdots + \mathbf{x}_r \otimes \mathbf{y}_r \otimes \mathbf{z}_r.$$

This reveals the true chemical factors responsible for the data: $r = \text{rank}(A)$ gives the number of pure substances in the mixtures, $\mathbf{x}_p = (x_{1p}, \ldots, x_{lp})$ gives the relative concentrations of pth substance in specimens $1, \ldots, l$; $\mathbf{y}_p = (y_{1p}, \ldots, y_{mp})$ gives the excitation spectrum of pth substance; and $\mathbf{z}_p = (z_{1p}, \ldots, z_{np})$ gives the emission spectrum of pth substance. The emission and excitation spectra would then allow one to identify the pure substances.

Of course, this is only valid in an idealized situation when measurements are performed perfectly without error and noise. Under realistic noisy circumstances, one would then need to a find best rank-r approximation, which is where the coherence results of Section VIII play a role. In this case, $\mu(\mathbf{x}_1, \ldots, \mathbf{x}_r)$ measures the relative abundance of the pure substances in the samples while $\mu(\mathbf{y}_1, \ldots, \mathbf{y}_r)$ and $\mu(\mathbf{z}_1, \ldots, \mathbf{z}_r)$ measure the spectroscopic likeness of these pure substances in the sense of absorbance and fluorescence respectively.

G. Statistical Independence Induces Diversity

We will now discuss a somewhat different way to achieve diversity. Assume the linear model below

$$\mathbf{x}(t) = U\mathbf{s}(t), \tag{68}$$

where only the signal $\mathbf{x}(t)$ is observed, $U = [\mathbf{u}_1, \ldots, \mathbf{u}_r]$ is an unknown $n \times r$ mixing matrix, and $\mathbf{s}(t) = (s_1(t), \ldots, s_r(t))$ has mutually statistically independent components. One may construct $K_d(\mathbf{x})$, the dth order cumulant hypermatrix [54] of $\mathbf{x}(t)$, and it will satisfy the multilinear model

$$K_d(\mathbf{x}) = \sum_{p=1}^{r} \lambda_p(\mathbf{s})\mathbf{u}_p \otimes \cdots \otimes \mathbf{u}_p$$

where $\lambda_p(\mathbf{s})$ denotes the pth diagonal entry of the dth cumulant hypermatrix of $\mathbf{s}$. Because of the statistical independence of $\mathbf{s}(t)$, the off-diagonal entries of the dth cumulant hypermatrix of $\mathbf{s}$ are zero [18], [54]. If $d \geq 3$, then the matrix U and the entries $\lambda_p(\mathbf{s})$ can be identified [18]. One may apply the results of Section VIII to deduce uniqueness of the solution.

Such problems generalize to convolutive mixtures and have applications in telecommunications, radar, sonar, speech processing, and biomedical engineering [18].

H. Nonstationarity Induces Diversity

If a signal $x(t)$ is nonstationary, its time-frequency transform, defined by

$$X(t, f) = \int x(u)\kappa(u - t; f)\, du$$

for some given kernel κ, bears information. If variables t and f are discretized, then the values of $X(t, f)$ can be stored in a matrix X; and the more nonstationary the signal $x(t)$, the larger the rank of X. A similar statement can be made on a signal $y(\mathbf{z})$ depending on a spatial variable $\mathbf{z}$. The discrete values of the space-wavevector transform $Y(\mathbf{z}, \mathbf{w})$ of a field $y(\mathbf{z})$ can be stored in a matrix Y; and the less homogeneous the field $y(\mathbf{z})$, the larger the rank of Y. This is probably the reason why algorithms proposed in [2], [73] permit one to localize and extract dipole contributions in the brains using a multilinear model, provided that one has distinct time-frequency or space-wavevector patterns. Nevertheless, such localization is guaranteed to be successful only under restrictive assumptions.

X. Further Work

A separate article discussing practical algorithms for the bounded coherence best rank-r multilinear approximation is under preparation with additional coauthors. These algorithms follow the general strategy of the greedy approximations WOGA and OGA discussed in Sections VIII-D and VIII-E but contain other elements exploiting the special separable structure of our problem. Extensive numerical experiments will be provided in the forthcoming article.

Acknowledgment

We thank Ignat Domanov for pointing out that in Theorem 25, the factor on the right hand side may be improved from r to $r - 1$, and that the improvement is sharp. We owe special thanks to Harm Derksen for pointing out an error in an earlier version of Section VII and for very helpful discussions regarding nuclear norm of tensors. Sections VIII-D and VIII-E came from an enlightening series of lectures Vladimir Temlyakov gave at the IMA in Minneapolis and the useful pointers he graciously provided afterwards. We gratefully acknowledge Tom Luo, Nikos Sidiropoulos, Yuan Yao, and two anonymous reviewers for their helpful comments.

References

[1] H. Abo, G. Ottaviani, and C. Peterson, "Induction for secant varieties of Segre varieties," *Trans. Amer. Math. Soc.*, vol. 361, no. 2, pp. 767–792, 2009.

[2] H. Becker, P. Comon, L. Albera, M. Haardt, and I. Merlet, "Multi-way space-time-wavevector analysis for EEG source separation," *Signal Process.*, vol. 92, no. 4, pp. 1021–1031, 2012.

[3] G. Beylkin and M. J. Mohlenkamp, "Numerical operator calculus in higher dimensions," *Proc. Nat. Acad. Sci.*, vol. 99, no. 16, pp. 10246–10251, 2002.

[4] D. Bini, G. Lotti, and F. Romani, "Approximate solutions for the bilinear form computational problem," *SIAM J. Comput.*, vol. 9, no. 4, pp. 692–697, 1980.

[5] C. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer-Verlag, 2006.

[6] G. Cain and G. H. Meyer, *Separation of Variables for Partial Differential Equations*. Cleveland, OH, USA: CRC Press, 2006.

[7] E. J. Candès and B. Recht, "Exact matrix completion via convex optimization," *Found. Comput. Math.*, vol. 9, no. 6, pp. 717–772, 2009.

[8] E. J. Candès and T. Tao, "The power of convex relaxation: Near-optimal matrix completion," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2053–2080, May 2010.

[9] E. J. Candès and J. Romberg, "Sparsity and incoherence in compressive sampling," *Inverse Problems*, vol. 23, no. 3, pp. 969–985, 2007.

[10] J. F. Cardoso and E. Moulines, "Asymptotic performance analysis of direction-finding algorithms based on fourth-order cumulants," *IEEE Trans. Signal Process.*, vol. 43, no. 1, pp. 214–224, Jan. 1995.

[11] J. D. Carroll and J. J. Chang,"Analysis of individual differences in multidimensional scaling via n-way generalization of Eckart-Young decomposition," *Psychometrika*, vol. 35, no. 3, pp. 283–319, 1970.

[12] M. V. Catalisano, A. V. Geramita, and A. Gimigliano, "Higher secant varieties of the Segre varieties," *J. Pure Appl. Algebra*, vol. 201, nos. 1–3, pp. 367–380, 2005.

[13] Y. Chen. (2012). *Inner Product Quadratures*, [Online]. Available: http://arxiv.org/abs/1205.0601

[14] L. Chiantini and C. Ciliberto, "On the concept of k-sectant order of a variety," *J. London Math. Soc.*, vol. 73, no. 2, pp. 436–454, 2006.

[15] A. Cohen, W. Dahmen, and R. DeVore, "Compressed sensing and best k-term approximation," *J. Amer. Math. Soc.*, vol. 22, no. 1, pp. 211–231, 2009.

[16] P. Comon, "Contrasts, independent component analysis, and blind deconvolution," *Internat. J. Adapt. Control Signal Process.*, vol. 18, no. 3, pp. 225–243, 2004.

[17] P. Comon, "Independent component analysis," in *Higher Order Statistics*, J.-L. Lacoume, Ed. New York, NY, USA: Elsevier, 1992, pp. 29–38.

[18] P. Comon and C. Jutten, *Handbook of Blind Source Separation, Independent Component Analysis and Applications*. San Diego, CA, USA: Academic Press, 2010.

[19] P. Comon, X. Luciani, and A. L. F. De Almeida, "Tensor decompositions, alternating least squares and other tales," *J. Chemometr.*, vol. 23, nos. 7–8, pp. 393–405, 2009.

[20] G. Cybenko, "Approximation by superpositions of sigmoidal functions," *Math. Control Signals Syst.*, vol. 2, no. 4, pp. 303–314, 1989.

[21] V. De Silva and L.-H. Lim, "Tensor rank and the ILL-posedness of the best low-rank approximation problem," *SIAM J. Matrix Anal. Appl.*, vol. 30, no. 3, pp. 1084–1127, 2008.

[22] A. Defant and K. Floret, *Tensor Norms and Operator Ideals*, vol. 176. The Netherlands: Amsterdam, 1993.

[23] J. Demmel, I. Dumitriu, and O. Holtz, "Fast linear algebra is stable," *Numer. Math.*, vol. 108, no. 1, pp. 59–91, 2007.

[24] H. Derksen, "Kruskal's uniqueness inequality is sharp," *Linear Algebra Appl.*, vol. 438, no. 2, pp. 708–712, 2013.

[25] H. Derksen, "On the nuclear norm and the singular value decomposition of tensors," *Numer. Anal.*, vol. 25, pp. 1–3, Aug. 2013.

[26] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[27] D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ^1 minimization," *Proc. Nat. Acad. Sci.*, vol. 100, no. 5, pp. 2197–2202, 2003.

[28] D. L. Donoho, M. Elad, and V. N. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 6–18, Jan. 2006.

[29] D. L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decompositions," *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 2845–2862, Nov. 2001.

[30] M. Fazel, H. Hindi, and S. Boyd, "A rank minimization heuristic with application to minimum order system approximation," in *Proc. Amer. Control Conf.*, Jun. 2001, pp. 4734–4739.

[31] C. E. R. Fernandes, G. Favier, and J. C. M. Mota, "Blind channel identification algorithms based on the PARAFAC decomposition of cumulant tensors: The single and multiuser cases," *Signal Process.*, vol. 88, no. 6, pp. 1381–1401, 2008.

[32] A. C. Gilbert, S. Muthukrishnan, and M. J. Strauss, "Approximation of functions over redundant dictionaries using coherence," in *Proc. SODA*, vol. 14. 2003, pp. 243–252.

[33] R. Gribonval and M. Nielsen, "Sparse representations in unions of bases," *IEEE Trans. Inf. Theory*, vol. 49, no. 12, pp. 3320–3325, Dec. 2003.

[34] D. Gross, "Recovering low-rank matrices from few coefficients in any basis," *IEEE Trans. Inf. Theory*, vol. 57, no. 3, pp. 1548–1566, Mar. 2011.

[35] X. Guo, S. Miron, D. Brie, S. Zhu, and X. Liao, "A CP perspective on uniqueness of DOA estimation using a vector sensor array," *IEEE Trans. Signal Process.*, vol. 59, no. 7, pp. 3475–3481, Jul. 2011.

[36] R. A. Harshman, "Foundations of the PARAFAC procedure: Models and conditions for an explanatory multimodal factor analysis," in *Proc. UCLA Working Papers Phonetics*, vol. 16, 1970, pp. 1–84.

[37] F. L. Hitchcock, "The expression of a tensor or a polyadic as a sum of products," *J. Math. Phys.*, vol. 6, no. 1, pp. 164–189, 1927.

[38] T. Jiang and N. Sidiropoulos, "Kruskal's permutation lemma and the identification of CANDECOMP PARAFAC and bilinear models," *IEEE Trans. Signal Process.*, vol. 52, no. 9, pp. 2625–2636, Sep. 2004.

[39] J.-P. Kahane, "Sur le théorème de superposition de Kolmogorov," *J. Approximation Theory*, vol. 13, no. 3, pp. 229–234, 1975.

[40] A. N. Kolmogorov, "On the representation of continuous functions of several variables by superpositions of continuous functions of a smaller number of variables," *Amer. Math. Soc. Trans.*, vol. 17, pp. 369–373, Feb. 1961.

[41] J. B. Kruskal, "Three-way arrays: Rank and uniqueness of trilinear decompositions," *Linear Algebra Appl.*, vol. 18, no. 2, pp. 95–138, 1977.

[42] J. Håstad, "Tensor rank is NP-complete," *J. Algorithms*, vol. 11, no. 4, pp. 644–654, 1990.

[43] C. J. Hillar and L.-H. Lim, "Most tensor problems are NP-hard," *J. ACM*, vol. 60, no. 6, Art. 45, pp. 1–39, 2013.

[44] J.-B. Hiriart-Urruty and C. Lemaréchal, *Convex Analysis and Minimization Algorithms*, vol. 305. New York, NY, USA: Springer-Verlag, 1993.

[45] C. Itzykson and J. M. Drouffe, *Statistical Field Theory*, Cambridge, U.K.: Cambridge Univ. Press, 1989.

[46] L.-H. Lim, "Singular values and eigenvalues of tensors: A variational approach," in *Proc. 1st IEEE Int. Workshop CAMSAP*, vol. 1. Dec. 2005, pp. 129–132.

[47] L.-H. Lim, "Tensors and hypermatrices," in *Handbook of Linear Algebra*, 2nd ed., L. Hogben Ed. Cleveland, OH, USA: CRC Press, 2013.

[48] L.-H. Lim and P. Comon, "Nonnegative approximations of nonnegative tensors," *J. Chemometr.*, vol. 23, nos. 7–8, pp. 432–441, 2009.

[49] L.-H. Lim and P. Comon, "Multiarray signal processing: Tensor decomposition meets compressed sensing," *Comptes Rrndus Mécanique*, vol. 338, no. 6, pp. 311–320, 2010.

[50] E. D. Livshitz, "On the optimality of the orthogonal greedy algorithm for μ-coherent dictionaries," *J. Approx. Theory*, vol. 164, no. 5, pp. 668–681, 2012.

[51] X. Luciani, S. Mounier, R. Redon, and A. Bois, "A simple correction method of inner filter effects affecting FEEM and its application to the PARAFAC decomposition," *Chemometr. Intell. Lab. Syst.*, vol. 96, no. 2, pp. 227–238, 2009.

[52] S. Mallat, *A Wavelet Tour of Signal Processing: The Sparse Way*, 3rd ed., San Diego, CA, USA: Academic Press, 2009.

[53] D. K. Maslen and D. N. Rockmore, "Separation of variables and the computation of Fourier transforms on finite groups," *J. Amer. Math. Soc.*, vol. 10, no. 1, pp. 169–214, 1997.

[54] P. Mccullagh, *Tensor Methods in Statistics*. London, U.K.: Chapman & Hall, 1987.

[55] W. Miller, *Symmetry and Separation of Variables*, vol. 4. Reading, MA, USA: Addison-Wesley, 1977.

[56] A. Nehorai and E. Paldi, "Vector-sensor array processing for electromagnetic source localization," *IEEE Trans. Signal Process.*, vol. 42, no. 2, pp. 376–398, Feb. 1994.

[57] P. Paatero, "A weighted non-negative least squares algorithm for three-way PARAFAC factor analysis," *Chemometr. Intell. Lab. Syst.*, vol. 38, no. 2, pp. 223–242, 1997.

[58] K. Pearson, "On lines and planes of closest fit to systems of points in space," *Phil. Mag.*, vol. 2, no. 11, pp. 559–572, 1901.

[59] J. C. Pesquet and E. Moreau, "Cumulant based independence measures for linear mixtures," *IEEE Trans. Inf. Theory*, vol. 47, no. 5, pp. 1947–1956, Jul. 2001.

[60] S. Shamsunder and G. B. Giannakis, "Modeling of non-Gaussian array data using cumulants: DOA estimation of more sources with less sensors," *Signal Process.*, vol. 30, no. 3, pp. 279–297, 1993.

[61] N. D. Sidiropoulos and R. Bro, "On the uniqueness of multilinear decomposition of N-way arrays," *J. Chemometr.*, vol. 14, no. 3, pp. 229–239, 2000.

[62] N. D. Sidiropoulos, R. Bro, and G. B. Giannakis, "Parallel factor analysis in sensor array processing," *IEEE Trans. Signal Process.*, vol. 48, no. 8, pp. 2377–2388, Aug. 2000.

[63] N. D. Sidiropoulos, G. B. Giannakis, and R. Bro, "Blind PARAFAC receivers for DS-CDMA systems," *IEEE Trans. Signal Process.*, vol. 48, no. 3, pp. 810–823, Mar. 2000.

[64] A. Smilde, R. Bro, and P. Geladi, *Multi-Way Analysis*. New York, NY, USA: Wiley, 2004.

[65] A. Stegeman and P. Comon, "Subtracting a best rank-1 approximation does not necessarily decrease tensor rank," *Linear Algebra Appl.*, vol. 433, no. 7, pp. 1276–1300, 2010.

[66] E. Stein and G. Weiss, *Harmonic Analysis in Euclidean Space*. Princeton, NJ, USA: Princeton Univ. Press, 1971.

[67] V. Strassen, "Gaussian elimination is not optimal," *Numer. Math.*, vol. 13, no. 4, pp. 354–356, 1969.

[68] V. N. Temlyakov, *Greedy Approximations*. Cambridge U.K.: Cambridge Univ. Press, 2011.

[69] V. N. Temlyakov, *Greedy Approximations*, vol. 331. Cambridge U.K.: Cambridge Univ. Press, 2006, pp. 371–394.

[70] A. Tillmann and M. E. Pfetsch. (2012). *The Computational Complexity of the Restricted Isometry Property, the Nullspace Property, and Related Concepts in Compressed Sensing* [Online]. Available: http://arxiv.org/abs/1205.2081

[71] M. E. Tipping and C. M. Bishop, "Probabilistic principal component analysis," *J. R. Stat. Soc. B*, vol. 61, no. 3, pp. 611–622, 1999.

[72] J. A. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.

[73] M. Weis, F. Roemer, M. Haardt, D. Jannek, and P. Husar, "Multi-dimensional space-time-frequency component analysis of event related EEG data using closed-form PARAFAC," in *Proc. IEEE ICASSP*, Apr. 2009, pp. 349–352.

[74] R. M. Young, *An Introduction to Nonharmonic Fourier Series*, 1st ed. San Diego, CA, USA: Academic Press, 2001.

Lek-Heng Lim received his undergraduate education at the National University of Singapore and his postgraduate education at Cornell, Cambridge, and Stanford Universities. He was a Charles Morrey assistant professor at the University of California, Berkeley from 2007–2010 and has been an assistant professor at the University of Chicago since 2010. He is a member of the editorial boards of *Linear Algebra and its Applications* and *Linear and Multilinear Algebra*.

Pierre Comon (M'87–SM'95–F'07) graduated in 1982, and received the Doctorate degree in 1985, both from the University of Grenoble, France. He later received the Habilitation to Lead Researches in 1995, from the University of Nice, France. He has been for nearly 13 years in industry, first with Crouzet-Sextant, Valence, France, between 1982 and 1985, and then with Thomson Marconi, Sophia-Antipolis, France, between 1988 and 1997. He spent 1987 with the ISL laboratory, Stanford University, CA. He joined in 1997 the Eurecom Institute, Sophia-Antipolis, France. He is research director with CNRS since 1998, first at laboratory I3S, Sophia-Antipolis, France, until 2012, and then at Gipsa-Lab, Grenoble, France. His research interests include High-Order Statistics (HOS), Blind techniques, Statistical Signal and Array Processing, Tensor decompositions, Multi-Way Factor Analysis and its applications to biomedical end environment.

Dr. Comon was Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 1995 to 1998, and a member of the French National Committee of Scientific Research from 1995 to 2000. He was the coordinator of the European Basic Research Working Group on HOS, ATHOS, from 1992 to 1995. Between 1992 and 1998, he was a member of the Technical and Scientific Council of the Thomson Group. Between 2001 and 2004 he acted as launching Associate Editor with the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS I, in the area of Blind Techniques. He has also been a member of the editorial board of the *Elsevier Journal Signal Processing*, and member of several IEEE TC. He is presently in the editorial board of *SIAM Journal on Matrix Analysis and Applications*.