Kruskal's Permutation Lemma and the Identification of CANDECOMP/PARAFAC and Bilinear Models with Constant Modulus Constraints

Tao Jiang, *Student Member, IEEE,* and Nicholas D. Sidiropoulos, *Senior Member, IEEE*

***Abstract*—CANDECOMP/PARAFAC (CP) analysis is an extension of low-rank matrix decomposition to higher-way arrays, which are also referred to as *tensors*. CP extends and unifies several array signal processing tools and has found applications ranging from multidimensional harmonic retrieval and angle-carrier estimation to blind multiuser detection. The uniqueness of CP decomposition is not fully understood yet, despite its theoretical and practical significance. Toward this end, we first revisit Kruskal's Permutation Lemma, which is a cornerstone result in the area, using an accessible basic linear algebra and induction approach. The new proof highlights the nature and limits of the identification process. We then derive two equivalent necessary and sufficient uniqueness conditions for the case where one of the component matrices involved in the decomposition is full column rank. These new conditions explain a curious example provided recently in a previous paper by Sidiropoulos, who showed that Kruskal's condition is in general sufficient but not necessary for uniqueness and that uniqueness depends on the particular joint pattern of zeros in the (possibly pretransformed) component matrices. As another interesting application of the Permutation Lemma, we derive a similar necessary and sufficient condition for unique bilinear factorization under constant modulus (CM) constraints, thus providing an interesting link to (and unification with) CP.**

***Index Terms*—CANDECOMP, constant modulus, identifiablity, PARAFAC, SVD, three-way array analysis, uniqueness.**

I. Introduction

LINEAR algebra plays an important role in modern signal processing, as evidenced by recent issues of this TRANSACTIONS. Various matrix decompositions are routinely used to prove results and as building blocks in the construction of signal processing algorithms. In many signal processing applications of linear algebra tools, the signal part of a postulated model lies in a so-called *signal subspace*, whereas the parameters of interest are in one-to-one correspondence with a certain basis of this subspace. The signal subspace can often be reliably estimated from measured data, but the particular basis of interest cannot be identified without additional problem-specific structure.

Consider a $I \times J$ matrix $\mathbf{X}$ of rank F. By definition of matrix rank, $\mathbf{X}$ can be decomposed as a sum of F rank one matrices

$$\mathbf{X} = \sum_{f}^{F} \mathbf{a}_f \mathbf{b}_f^T. \tag{1}$$

Let $\mathbf{a}_f$ ($\mathbf{b}_f$) be the fth column of $\mathbf{A}$ (resp. $\mathbf{B}$). $\mathbf{X}$ can then be written as $\mathbf{X} = \mathbf{A}\mathbf{B}^T$. An equivalent scalar view of (1) is given by the *bilinear decomposition*

$$x_{i,j} = \sum_{f}^{F} a_{i,f} b_{j,f} \tag{2}$$

where $x_{i,j}$ ($a_{i,f}, b_{j,f}$) denotes the (i,j)th (resp. (i,f), (j,f)) entry of $\mathbf{X}$ (resp. $\mathbf{A}$, $\mathbf{B}$). Clearly, the decomposition in (1) [or, equivalently, (2)] is not unique because $\mathbf{X} = \mathbf{A}\mathbf{B}^T = \mathbf{A}\mathbf{M}\mathbf{M}^{-1}\mathbf{B}^T$ for any invertible $\mathbf{M}$. The only notable exception is when $\mathbf{X}$ is of rank one, in which case, only the trivial scaling ambiguity remains. In applications, the typical way around nonuniqueness is the imposition of application-specific structural properties on the matrix factors. Common examples include orthogonality (as in SVD), Vandermonde, Toeplitz, constant modulus (CM), or finite-alphabet (FA) constraints.

In sharp contrast to the case of matrices (also known as *two-way* arrays because they are indexed by two independent variables), low-rank decomposition of three-way arrays (also known as *tensors*) is unique under certain relatively mild conditions [8]. Let $\underline{\mathbf{X}}$ be a rank-F three way array of order $I \times J \times K$ with typical element $x_{i,j,k}$, and consider the F component trilinear decomposition

$$x_{i,j,k} = \sum_{f}^{F} a_{i,f} b_{j,f} c_{k,f}. \tag{3}$$

Let $a_{i,f}$ ($b_{j,f}, c_{k,f}$) be the typical element of matrix $\mathbf{A}$ (resp. $\mathbf{B}$, $\mathbf{C}$). It has been shown by Kruskal [8] that, under certain conditions, the component matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ can be uniquely identified from $\underline{\mathbf{X}}$ up to column permutation and scaling.

The model in (3) was independently introduced in 1970 by two different groups as CANonical DECOMPosition [3] (CANDECOMP) and PARAllel FACtor analysis [4](PARAFAC),

Manuscript received June 6, 2003; revised October 1, 2003. This work was supported by the Army Research Laboratory through participation in the ARL Collaborative Technology Alliance (ARL-CTA) for Communications and Networks under Cooperative Agreement DADD19-01-2-0011 and the National Science Foundation under Grants NSF/CAREER CCR-0096165 and NSF/Wireless IT & Networks CCR-0096164. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Martin Haardt.

T. Jiang is with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: jiang@ece.umn.edu).

N. D. Sidiropoulos is with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA, and also with the Technical University of Crete, Chania-Crete 73100, Greece (e-mail: nikos@ece.umn.edu).

Digital Object Identifier 10.1109/TSP.2004.832022

respectively. Nowadays, the term CANDECOMP/PARAFAC (CP) is often used and will be adopted in this paper. The CP model is a cornerstone of *three-way analysis*, which deals with algebraic and numerical aspects of tensorial decomposition. Three-way analysis was first introduced in psychometrics and has gradually found many applications in diverse disciplines, including arithmetic complexity, statistics, chemometrics, and, more recently, signal processing for communications.

Three-way methods are often naturally applicable for the analysis of multidimensional data sets encountered in blind identification [13], multiuser signal separation [6], [11], [14], and diversity systems [7]. For example, in the context of uplink reception for narrowband cellular DS-CDMA systems with symbol-periodic spreading and a base station antenna array, the received baseband-equivalent data constitutes a three-way data array indexed by antenna element, symbol epoch, and chip. The signal part of this data array can be modeled by a three-way array [as in (3)] of rank equal to the number of users in the system [14]. In this application of CP, $a_{i,f}$ corresponds to the channel gain between user f and antenna i, $b_{j,f}$ is the jth symbol of user f, and $c_{k,f}$ is the kth chip of user f. Likewise, I stands for the number of receive antennas, J for the number of symbol snapshots, and K for the spreading gain. In this context, the uniqueness properties of CP imply, e.g., that blind identification is possible with more users than spreading, antennas, and symbols simultaneously [14]. Another interesting signal processing application of CP appears in the context of multiple invariance antenna array processing. Therein, $c_{k,f}$ is the displacement-induced phase shift for user f when "hopping" from the ith element of the reference subarray to the corresponding element in the kth displaced but otherwise identical subarray. In this context, the application of CP ideas pinned down a long-standing identifiability issue [13].

The usefulness of the CP model is mainly due to its ubiquitous uniqueness properties and its direct link to low-rank decomposition. The uniqueness of CP decomposition enables unambiguous interpretation of the estimated model parameters, which is crucial in many applications. Theoretical interests aside, deeper understanding of CP uniqueness issues could pave the way for more general algebraic decomposition and improved iterative fitting algorithms.

The in-depth study of uniqueness of CP decomposition entails technical challenges. In seminal work back in 1977 [8], Kruskal proved that uniqueness is guaranteed, provided that the sum of k-ranks of the three component matrices[1] is no less than twice the rank of the CP model plus 2. In [16], ten Berge and Sidiropoulos have shown that Kruskal's sufficient condition is *not* necessary in general and that uniqueness may still hold even if Kruskal's condition is violated. Surprisingly, they have also found out via simple examples that the uniqueness of CP decomposition depends on the particular joint pattern of zeros in the (possibly pretransformed) component matrices. These curious examples may lead one to believe that pursuing a necessary and sufficient condition for uniqueness of CP decomposition is probably hopeless.

[1]For a definition of k-rank, see subsection A (Notation) of this Section.

This unusual phenomenon will be understood in this paper. In fact, when one of the component matrices involved in the decomposition is full column rank, two equivalent necessary and sufficient uniqueness conditions are derived herein. One helps us understand the nature of identification in very intuitive terms. The other provides the means to check if a given decomposition is unique. With the aid of these necessary and sufficient uniqueness conditions, the explanation of the seemingly mysterious examples in [16] is straightforward.

Assuming that at least one of $\mathbf{A}, \mathbf{B}$, or $\mathbf{C}$ is full column rank is typically not restrictive in applications, usually, the sample size and distribution along at least one dimension are adequate to guarantee this with very high probability. The methodology developed herein can be also extended to general CP models and offers the possibility to examine uniqueness of CP solutions, albeit the associated necessary and sufficient uniqueness conditions appear too complicated to verify in practice. We will see that the uniqueness of CP decomposition is "nonlinear" in nature,[2] whereas Kruskal's condition is probably the best "linear" sufficient condition for uniqueness.

In addition to algebraic structure, communications signals often exhibit FA or CM properties. As mentioned earlier, although bilinear decomposition is not unique in general, bilinear decomposition under FA/CM constraints can be unique [9], [15], [17]. We show that the necessary and sufficient condition for unique bilinear decomposition under CM constraints is strikingly similar to the one for uniqueness of certain CP models. The impact of additional diversity dimensions (i.e., three-way data) on FA/CM separation has yet to be addressed in the literature; all previous works dealt with a bilinear problem. We therefore also consider uniqueness for CP models under CM constraints. These developments open a path for cross-fertilization and unification of the literature on CP and CM models.

The rest of this paper is structured as follows. Section II lays out the model and problem statement. Section III provides a starting point and road map for proving CP uniqueness. The main new results are presented in Section IV. Explanation of the mysterious examples in [16] is given in Section V, which also establishes the promised link between the conditions for unique decomposition of certain restricted CP models on one hand and bilinear models subject to CM constraints on the other. In particular, we focus on CP models with at least one of the three component matrices being full column rank. We use the term *restricted CP* to refer to these models for brevity, but note again that one matrix being full column rank is typically not restrictive in practice. A new self-contained and intuitive proof of Kruskal's Permutation Lemma can be found in the Appendix . Conclusions are drawn in Section VI.

A. Notation

^{H} is the Hermitian (conjugate) transpose,[3] and $\omega(\mathbf{x})$ is the number of nonzero elements of $\mathbf{x}$. Given matrices $\mathbf{A}$ $(I \times F)$

[2]Uniqueness cannot be assessed by individual properties of the matrices $\mathbf{A}, \mathbf{B}$, or $\mathbf{C}$ but only by joint properties of the ensemble $(\mathbf{A}, \mathbf{B}, \mathbf{C})$.

[3]For brevity, we use the same notation for Hermitian transposition of matrices (and vectors) and conjugation of scalar quantities.

and $\mathbf{B}$ $(J \times F)$, their Khatri–Rao product $\mathbf{A} \odot \mathbf{B}$ $(JI \times F)$ is the matrix whose fth column is the Kronecker product of the respective columns of $\mathbf{A}$ and $\mathbf{B}$. $r_{\mathbf{A}}$ is the rank of $\mathbf{A}$; $k_{\mathbf{A}}$ is the Kruskal-rank of $\mathbf{A}$, the maximum number of linearly independent columns that can be drawn from $\mathbf{A}$ in an arbitrary fashion; $|\mathbf{A}|$ is the determinant of matrix $\mathbf{A}$; $diag(\mathbf{x})$ is a diagonal matrix containing the elements of vector $\mathbf{x}$. Among the various equivalent definitions of the rank of a matrix $\mathbf{A}$, we pay particular attention to the one given by Sylvester in 1851 (e.g., [2, p. 215]): *The rank of* $\mathbf{A}$ *is equal to the maximum order of nonzero minors of* $\mathbf{A}$. *A kth-order minor of* $\mathbf{A}$ *is the determinant of a* $k \times k$ *submatrix of* $\mathbf{A}$. Throughout, $Null(\mathbf{A})$ means left null, i.e., $Null(\mathbf{A}) := \{ \mathbf{x} \mid \mathbf{x}^H\mathbf{A} = \mathbf{0}\}$. $Span(\mathbf{A})$ denotes the linear space spanned by the columns of $\mathbf{A}$. In the Appendix , we will make extensive use of the equivalence

$$Span(\mathbf{A}) \subseteq Span(\bar{\mathbf{A}}) \Leftrightarrow Null(\mathbf{A}) \supseteq Null(\bar{\mathbf{A}}).$$

II. CANDECOMP/PARAFAC (CP) MODEL AND PROBLEM STATEMENT

Let $\underline{\mathbf{X}}$ be a rank-F three-way array of order $I \times J \times K$. The CP model decomposes $\underline{\mathbf{X}}$ into a sum of F triads (rank-one three-way factors), as in (3), which is reproduced here for ease of exposition:

$$x_{i,j,k} = \sum_{f}^{F} a_{i,f} b_{j,f} c_{k,f}. \tag{4}$$

The focus of this paper is the identification of the CP model parameters $a_{i,f}, b_{j,f}, c_{k,f}$ from the data $x_{i,j,k}$. More specifically, we are interested in conditions under which these parameters are identifiable from the data.

In the signal processing community, it is customary to ignore noise-induced model errors when studying identifiability. The reason is that identifiability is mostly appreciated as a sanity check for the high signal-to-noise ratio (SNR) regime. A different viewpoint is often taken in the data analysis community. Given a decomposition (obtained, e.g., through alternating least squares optimization [4], [14] in the case of CP)

$$y_{i,j,k} = \sum_{f}^{F} a_{i,f} b_{j,f} c_{k,f} + e_{i,j,k}$$

where $e_{i,j,k}$ denotes modeling errors, and bringing the error term to the left

$$y_{i,j,k} - e_{i,j,k} = \sum_{f}^{F} a_{i,f} b_{j,f} c_{k,f}$$

the following question is often of interest: *When is the signal part of the fitted model unique (conditioned on model errors)?* Both viewpoints are valid and boil down to the same mathematical problem statement.

Define an $I \times F$ matrix $\mathbf{A}$ with typical element $\mathbf{A}(i,f) = a_{i,f}$, $J \times F$ matrix $\mathbf{B}$ with $\mathbf{B}(j,f) = b_{j,f}$, $K \times F$ matrix $\mathbf{C}$ with $\mathbf{C}(k,f) = c_{k,f}$, and $J \times K$ matrices $\mathbf{X}_i$ with $\mathbf{X}_i(j,k) = x_{i,j,k}$. The model in (4) can be written as

$$\mathbf{X}_i = \mathbf{B}\, diag(\mathbf{a}_i^T)\mathbf{C}^T \tag{5}$$

where $\mathbf{a}_i^T$ stands for the ith row of $\mathbf{A}$. If we stack the $\mathbf{X}_i$ one over another, a compact matrix representation of the model in (4) is possible by employing the Khatri–Rao (column-wise Kronecker) product

$$\mathbf{X}^{JI\times K} := \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \vdots \\ \mathbf{X}_I \end{bmatrix} = \begin{bmatrix} \mathbf{B}\, diag(\mathbf{a}_1^T) \\ \mathbf{B}\, diag(\mathbf{a}_2^T) \\ \vdots \\ \mathbf{B}\, diag(\mathbf{a}_I^T)) \end{bmatrix} \mathbf{C}^T = (\mathbf{A} \odot \mathbf{B})\mathbf{C}^T. \tag{6}$$

By symmetry, $\mathbf{A}, \mathbf{B}, \mathbf{C}$ may switch their places in (6) if the modes of the array are switched accordingly.

Suppose we have two different decompositions of the same array $\underline{\mathbf{X}}$, namely, there is a triple $(\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}})$ other than $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ such that $(\mathbf{A} \odot \mathbf{B})\mathbf{C}^T = (\bar{\mathbf{A}} \odot \bar{\mathbf{B}})\bar{\mathbf{C}}^T$. Note that if $\mathbf{\Pi}$ is a permutation matrix and $\mathbf{\Lambda_A}, \mathbf{\Lambda_B}, \mathbf{\Lambda_C}$ are diagonal matrices such that $\mathbf{\Lambda_A}\mathbf{\Lambda_B}\mathbf{\Lambda_C} = \mathbf{I}$, then $(\mathbf{A\Pi\Lambda_A}, \mathbf{B\Pi\Lambda_B}, \mathbf{C\Pi\Lambda_C})$ will yield the same array given by $(\mathbf{A}, \mathbf{B}, \mathbf{C})$. A CP decomposition of the data in (6) is therefore said to be unique if for every other decomposition of (6), which is $\mathbf{X}^{JI\times K} = (\bar{\mathbf{A}} \odot \bar{\mathbf{B}})\bar{\mathbf{C}}^T$, we have $\bar{\mathbf{A}} = \mathbf{A\Pi\Lambda_A}, \bar{\mathbf{B}} = \mathbf{B\Pi\Lambda_B}, \bar{\mathbf{C}} = \mathbf{C\Pi\Lambda_C}$ for some permutation matrix $\mathbf{\Pi}$ and diagonal matrices $\mathbf{\Lambda_A}, \mathbf{\Lambda_B}, \mathbf{\Lambda_C}$ with $\mathbf{\Lambda_A}\mathbf{\Lambda_B}\mathbf{\Lambda_C} = \mathbf{I}$.

Kruskal has shown [8] that if $k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}} \geq 2F + 2$, then CP decomposition is unique. A conjecture that Kruskal's condition is necessary and sufficient has been upheld until the recent work [16]. An approach to finding alternative CP decompositions for any given CP model has been developed in [16] to study the uniqueness of CP decomposition, and Kruskal's conjecture was failed by simple counterexample. However, [16] did not further qualify the uniqueness conditions. In what follows, we will derive two equivalent necessary and sufficient uniqueness conditions when one of the component matrices involved in the decomposition is full column rank and explain the examples in [16]. As a bonus, we will establish a link to uniqueness of bilinear factorization under CM constraints. Last, but not least, we will provide a more palatable proof of the Permutation Lemma, which, at least, we would have appreciated being readily available several years ago.

III. ROADMAP OF UNIQUENESS RESULTS

Recall that CP decomposition, when unique, is unique up to a common permutation and nonsingular scaling/counter-scaling of columns of the component matrices. In hindsight, it is therefore natural to ask under what conditions two matrices are the same up to permutation and scaling of columns. This is precisely the subject of Kruskal's Permutation Lemma [8]:

Lemma 1: [8] Suppose we are given two matrices $\mathbf{A}$ and $\bar{\mathbf{A}}$, which are $I \times F$ and $I \times \bar{F}$. Suppose $\mathbf{A}$ has no zero columns. If for any vector $\mathbf{x} \in \mathbb{C}^N$ such that

$$\omega(\mathbf{x}^H\bar{\mathbf{A}}) \leq F - r_{\bar{\mathbf{A}}} + 1 \tag{7}$$

we have

$$\omega(\mathbf{x}^H\mathbf{A}) \leq \omega(\mathbf{x}^H\bar{\mathbf{A}})$$

then $F \leq \bar{F}$; if also $F \geq \bar{F}$, then $F = \bar{F}$, and there exist a permutation matrix $\mathbf{P}_{\bar{\mathbf{A}}}$ and a nonsingular diagonal matrix $\mathbf{\Lambda}$ such that $\mathbf{A} = \bar{\mathbf{A}}\mathbf{P}_{\bar{\mathbf{A}}}\mathbf{\Lambda}$.

This lemma is *the* key tool in the area of CP analysis and the cornerstone for Kruskal's proof of uniqueness of CP decomposition. Since the statement reads as a sufficient condition, it is tempting to attempt to improve on Kruskal's condition for uniqueness of CP decomposition by sharpening the condition in the permutation lemma. However, Kruskal's proof of the permutation lemma is ingenious but also largely inaccessible. We managed to reprove Kruskal's Permutation Lemma using a systematic basic linear algebra and induction approach (see the Appendix). The new proof suggests that the condition in Kruskal's Permutation Lemma is sharp; hence, the aforementioned attempt is unlikely to succeed.

Necessary conditions for CP uniqueness are worth recounting at this point. One is that neither $\mathbf{A}$, nor $\mathbf{B}$, nor $\mathbf{C}$ has a pair of proportional columns [5]. Another is that the Khatri–Rao product of any two component matrices must be full column rank [10].

In hindsight, the proof for uniqueness of CP decomposition can be decoupled into three separate steps. Given $(\mathbf{A}\odot\mathbf{B})\mathbf{C}^T = (\bar{\mathbf{A}}\odot\bar{\mathbf{B}})\bar{\mathbf{C}}^T$, the first step is to show that $\bar{\mathbf{A}} = \mathbf{A}\mathbf{\Pi}_{\mathbf{A}}\mathbf{\Lambda}_{\mathbf{A}}, \bar{\mathbf{B}} = \mathbf{B}\mathbf{\Pi}_{\mathbf{B}}\mathbf{\Lambda}_{\mathbf{B}}$, the second step is to show that $\mathbf{\Pi}_{\mathbf{A}} = \mathbf{\Pi}_{\mathbf{B}} = \mathbf{\Pi}$, and the last step is to show $\bar{\mathbf{C}} = \mathbf{C}\mathbf{\Pi}(\mathbf{\Lambda}_{\mathbf{A}}\mathbf{\Lambda}_{\mathbf{B}})^{-1} = \mathbf{C}\mathbf{\Pi}\mathbf{\Lambda}_{\mathbf{C}}$. This last step is straightforward once the previous steps are finished:

$$\begin{aligned}(\mathbf{A}\odot\mathbf{B})\mathbf{C}^T &= (\bar{\mathbf{A}}\odot\bar{\mathbf{B}})\bar{\mathbf{C}}^T = ((\mathbf{A}\mathbf{\Pi}\mathbf{\Lambda}_{\mathbf{A}})\odot(\mathbf{B}\mathbf{\Pi}\mathbf{\Lambda}_{\mathbf{B}}))\bar{\mathbf{C}}^T \\ &= (\mathbf{A}\odot\mathbf{B})\mathbf{\Pi}\mathbf{\Lambda}_{\mathbf{A}}\mathbf{\Lambda}_{\mathbf{B}}\bar{\mathbf{C}}^T\end{aligned}$$

and since $\mathbf{A}\odot\mathbf{B}$ is full column rank (recall that this is one of the necessary conditions for uniqueness), we have

$$\mathbf{C}^T = \mathbf{\Pi}\mathbf{\Lambda}_{\mathbf{A}}\mathbf{\Lambda}_{\mathbf{B}}\bar{\mathbf{C}}^T$$

or

$$\bar{\mathbf{C}} = \mathbf{C}\mathbf{\Pi}(\mathbf{\Lambda}_{\mathbf{A}}\mathbf{\Lambda}_{\mathbf{B}})^{-1}.$$

When one of the component matrices, say $\mathbf{C}$, is full column rank, the aforementioned procedure can be further simplified. One can first show that $\bar{\mathbf{C}} = \mathbf{C}\mathbf{\Pi}\mathbf{\Lambda}_{\mathbf{C}}$ and then obtain

$$(\mathbf{A}\odot\mathbf{B})\mathbf{C}^T = (\bar{\mathbf{A}}\odot\bar{\mathbf{B}})\bar{\mathbf{C}}^T = (\bar{\mathbf{A}}\odot\bar{\mathbf{B}})\mathbf{\Lambda}_{\mathbf{C}}\mathbf{\Pi}^T\mathbf{C}^T$$

and since $\mathbf{C}$ is full column rank

$$\bar{\mathbf{A}}\odot\bar{\mathbf{B}} = (\mathbf{A}\odot\mathbf{B})\mathbf{\Pi}\mathbf{\Lambda}_{\mathbf{C}}^{-1}$$

it then follows that $\bar{\mathbf{A}} = \mathbf{A}\mathbf{\Pi}\mathbf{\Lambda}_{\mathbf{A}}, \bar{\mathbf{B}} = \mathbf{B}\mathbf{\Pi}\mathbf{\Lambda}_{\mathbf{B}}$ for some $\mathbf{\Lambda}_{\mathbf{A}}$ and $\mathbf{\Lambda}_{\mathbf{B}}$, such that $\mathbf{\Lambda}_{\mathbf{A}}\mathbf{\Lambda}_{\mathbf{B}}\mathbf{\Lambda}_{\mathbf{C}} = \mathbf{I}$. Therefore, when $\mathbf{C}$ is full column rank, showing $\bar{\mathbf{C}} = \mathbf{C}\mathbf{\Pi}\mathbf{\Lambda}_{\mathbf{C}}$ is the key step. In Section IV, we will derive conditions under which this step can be accomplished. When those conditions do not hold, alternative CP decompositions will be constructed.

IV. Main Results

We now focus on proving uniqueness for restricted CP models, meaning those with full column rank $\mathbf{C}$. As we have seen in the previous section, one way to show that the decomposition of restricted CP models is unique is to prove that the full column rank component matrix $\mathbf{C}$ is unique up to permutation and column scaling. This entails conditions on both $\mathbf{A}$ and $\mathbf{B}$. $k_{\mathbf{A}} + k_{\mathbf{B}} \geq F + 2$, as a special case of Kruskal's conditions [8], can achieve the goal, but as shown in [16], this condition is not necessary.

The following condition will be proven to be necessary and sufficient to show that $\mathbf{C}$ is unique up to permutation and column scaling:

Condition A: None of the nontrivial linear combinations of columns of $\mathbf{A}\odot\mathbf{B}$ can be written as a tensor product of two vectors.

By *nontrivial* linear combinations, we mean those involving *at least two* columns of $\mathbf{A}\odot\mathbf{B}$.

Clearly, Condition A implies that $\mathbf{A}\odot\mathbf{B}$ is full column rank, since if $\mathbf{A}\odot\mathbf{B}$ is rank deficient, a nontrivial linear combination of columns of $\mathbf{A}\odot\mathbf{B}$ would constitute a zero vector, and this zero vector can be given in the form of tensor product of a zero vector and another vector.

Condition A also implies that neither $\mathbf{A}$ nor $\mathbf{B}$ has a pair of proportional columns. Otherwise, one can arrange a nontrivial linear combination of columns of $\mathbf{A}\odot\mathbf{B}$ such that the resulting vector is in the form of tensor product of two vectors.

Now, let us see why this condition is sufficient for the identification of restricted CP models. Suppose we have another decomposition of the same array $\underline{\mathbf{X}}$, $(\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}})$, such that $(\mathbf{A}\odot\mathbf{B})\mathbf{C}^T = (\bar{\mathbf{A}}\odot\bar{\mathbf{B}})\bar{\mathbf{C}}^T$. Thanks to the nonsingularity of $\mathbf{A}\odot\mathbf{B}$ implied by Condition A, it can be seen that $\mathbf{C}^T\mathbf{x} = \mathbf{0}$ for all $\mathbf{x}$ such that $\bar{\mathbf{C}}^T\mathbf{x} = \mathbf{0}$. This implies that $r_{\mathbf{C}} \leq r_{\bar{\mathbf{C}}}$. Since $\mathbf{C}$ is assumed full column rank, $\bar{\mathbf{C}}$ has to be full column rank as well.

To further proceed to show that $\bar{\mathbf{C}}$ is the same as $\mathbf{C}$ up to permutation and column scaling, we resort to Kruskal's Lemma. It suffices to show that $\omega(\mathbf{x}^H\mathbf{C}) \leq \omega(\mathbf{x}^H\bar{\mathbf{C}})$ for all $\omega(\mathbf{x}^H\bar{\mathbf{C}}) \leq 1$. Clearly, we only need to verify that $\omega(\mathbf{x}^H\mathbf{C}) \leq \omega(\mathbf{x}^H\bar{\mathbf{C}})$ holds for all $\omega(\mathbf{x}^H\bar{\mathbf{C}}) = 1$.

From $(\mathbf{A}\odot\mathbf{B})\mathbf{C}^T = (\bar{\mathbf{A}}\odot\bar{\mathbf{B}})\bar{\mathbf{C}}^T$, it follows that

$$(\bar{\mathbf{A}}\odot\bar{\mathbf{B}})\bar{\mathbf{C}}^T(\mathbf{x}^H)^T = (\mathbf{A}\odot\mathbf{B})\mathbf{C}^T(\mathbf{x}^H)^T, \quad \forall\mathbf{x}.$$

When $\mathbf{x}$ is such that $\omega(\mathbf{x}^H\bar{\mathbf{C}}) = 1$, $(\bar{\mathbf{A}}\odot\bar{\mathbf{B}})\bar{\mathbf{C}}^T(\mathbf{x}^H)^T$ is nothing but a scaled tensor product of a column of $\bar{\mathbf{A}}$ and the corresponding column of $\bar{\mathbf{B}}$. Therefore, $\omega(\mathbf{x}^H\mathbf{C})$ must be less than or equal to 1; otherwise, Condition A will be violated. Invoking Lemma 1, $\mathbf{C}$ and $\bar{\mathbf{C}}$ are the same up to permutation and column scaling. The result, therefore, follows.

To show necessity, we proceed by contradiction. Without loss of generality, we assume that $\mathbf{C}$ is an identity matrix,[4] $\mathbf{I}_F$, $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_F]$, $\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_F]$, and a linear combination of

[4]Under our working assumption of full column rank $\mathbf{C}$, this is without loss of generality in so far as uniqueness is concerned. This has been shown by ten Berge via suitable pretransformation of the data; see, e.g., [16].

the first two columns of $\mathbf{A} \odot \mathbf{B}$ constitutes a vector in the form of a tensor product of $\bar{\mathbf{a}}_1$ and $\bar{\mathbf{b}}_1$, i.e.,

$$\mathbf{a}_1 \otimes \mathbf{b}_1 + \mathbf{a}_2 \otimes \mathbf{b}_2 = \bar{\mathbf{a}}_1 \otimes \bar{\mathbf{b}}_1.$$

It is easy to see that

$$\begin{aligned}(\mathbf{A} \odot \mathbf{B})\mathbf{C}^T &= [\mathbf{a}_1 \otimes \mathbf{b}_1, \mathbf{a}_2 \otimes \mathbf{b}_2, \ldots, \mathbf{a}_F \otimes \mathbf{b}_F]\mathbf{I}_F \\ &= [\bar{\mathbf{a}}_1 \otimes \bar{\mathbf{b}}_1, \mathbf{a}_2 \otimes \mathbf{b}_2, \ldots, \mathbf{a}_F \otimes \mathbf{b}_F] \begin{bmatrix} 1 & 0 & \mathbf{0} \\ -1 & 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{F-2} \end{bmatrix} \\ &= (\bar{\mathbf{A}} \odot \bar{\mathbf{B}})\bar{\mathbf{C}}^T \end{aligned} \tag{8}$$

where

$$\begin{aligned}\bar{\mathbf{A}} &= [\bar{\mathbf{a}}_1, \mathbf{a}_2, \ldots, \mathbf{a}_F] \\ \bar{\mathbf{B}} &= [\bar{\mathbf{b}}_1, \mathbf{b}_2, \ldots, \mathbf{b}_F] \\ \bar{\mathbf{C}} &= \begin{bmatrix} 1 & 0 & \mathbf{0} \\ -1 & 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{F-2} \end{bmatrix}^T. \end{aligned} \tag{9}$$

Hence, $(\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}})$ constitutes an alternative decomposition. This completes the necessity part for Condition A.

Although Condition A has helped us intuitively understand the nature of the identification of restricted CP models, it has two limitations. The first is that Condition A is not easily verifiable. Second, and more important, is that the techniques used in the proof do not readily generalize to general CP models (rank-deficient $\mathbf{C}$). In the following, we will derive an alternative equivalent condition that is often better suited for verification and can be extended to cover general CP models.

We first define a set of $F \times F$ symmetric matrices $\mathbf{W}_{i_1,i_2,j_1,j_2}$ determined by the second-order minors of $\mathbf{A}$ and $\mathbf{B}$ as follows:

$$\mathbf{W}_{i_1,i_2,j_1,j_2}(f_1, f_2) = \frac{1}{2} \begin{vmatrix} a_{i_1,f_1} & a_{i_1,f_2} \\ a_{i_2,f_1} & a_{i_2,f_2} \end{vmatrix} \begin{vmatrix} b_{j_1,f_1} & b_{j_1,f_2} \\ b_{j_2,f_1} & b_{j_2,f_2} \end{vmatrix} \tag{10}$$

for $i_1 = 1, \ldots, I$, $i_2 = 1, \ldots, I$, $j_1 = 1, \ldots, J$, and $j_2 = 1, \ldots, J$.

We are now ready to derive the equivalent condition for identification of restricted CP models. As discussed before, it suffices that $\omega(\mathbf{x}^H\mathbf{C}) \le \omega(\mathbf{x}^H\bar{\mathbf{C}})$ for all $\omega(\mathbf{x}^H\bar{\mathbf{C}}) \le 1$. In particular, $\omega(\mathbf{x}^H\mathbf{C}) \le \omega(\mathbf{x}^H\bar{\mathbf{C}})$ for $\omega(\mathbf{x}^H\bar{\mathbf{C}}) = 1$.

Since

$$(\bar{\mathbf{A}} \odot \bar{\mathbf{B}})\bar{\mathbf{C}}^T(\mathbf{x}^H)^T = (\mathbf{A} \odot \mathbf{B})\mathbf{C}^T(\mathbf{x}^H)^T$$

invoking the identity [1] $vec(\mathbf{A}\,diag(\mathbf{x}^T)\mathbf{B}^T) = (\mathbf{A} \odot \mathbf{B})\mathbf{x}$, we have

$$\mathbf{A}\,diag(\mathbf{x}^H\mathbf{C})\mathbf{B}^T = \bar{\mathbf{A}}\,diag(\mathbf{x}^H\bar{\mathbf{C}})\bar{\mathbf{B}}^T.$$

Since $\omega(\mathbf{x}^H\bar{\mathbf{C}}) = 1$, we know

$$\mathbf{A}\,diag(\mathbf{x}^H\mathbf{C})\mathbf{B}^T = k_{f_0}\bar{\mathbf{a}}_{f_0}\bar{\mathbf{b}}_{f_0}^T \tag{11}$$

for some $f_0 \in \{1, \ldots, F\}$, where k_{f_0} is a nonzero constant, and therefore

$$r_{\mathbf{A}diag(\mathbf{x}^H\mathbf{C})\mathbf{B}^T} \le 1$$

which is equivalent to all the second-order minors of $\mathbf{A}\,diag(\mathbf{x}^H\mathbf{C})\mathbf{B}^T$ being zero.

Let $[y_1, \ldots, y_F] := \mathbf{x}^H\mathbf{C}$; all second-order minors of

$$\mathbf{M} := \mathbf{A}\,diag(\mathbf{x}^H\mathbf{C})\mathbf{B}^T = \sum_{f=1}^{F} y_f \mathbf{a}_f \mathbf{b}_f^T$$

must be equal to 0; equivalently

$$\begin{vmatrix} m_{i_1,j_1} & m_{i_1,j_2} \\ m_{i_2,j_1} & m_{i_2,j_2} \end{vmatrix} = \begin{vmatrix} \sum_{f=1}^{F} y_f a_{i_1,f} b_{j_1,f} & \sum_{f=1}^{F} y_f a_{i_1,f} b_{j_2,f} \\ \sum_{f=1}^{F} y_f a_{i_2,f} b_{j_1,f} & \sum_{f=1}^{F} y_f a_{i_2,f} b_{j_2,f} \end{vmatrix} = 0 \tag{12}$$

for $i_1 = 1, \ldots, I$, $i_2 = 1, \ldots, I$, and $j_1 = 1, \ldots, J$, $j_2 = 1, \ldots, J$.

Equation (12) can be written as

$$\left(\sum_{f=1}^{F} y_f a_{i_1,f} b_{j_1,f}\right)\left(\sum_{f=1}^{F} y_f a_{i_2,f} b_{j_2,f}\right) - \left(\sum_{f=1}^{F} y_f a_{i_1,f} b_{j_2,f}\right)\left(\sum_{f=1}^{F} y_f a_{i_2,f} b_{j_1,f}\right) = 0$$

which is nothing but

$$\sum_{h=1}^{F} \sum_{g \ne h} y_g y_h (a_{i_1,g} a_{i_2,h} b_{j_1,g} b_{j_2,h} - a_{i_1,g} a_{i_2,h} b_{j_1,h} b_{j_2,g}) = 0. \tag{13}$$

Further simplifying (13), we obtain (14), shown at the bottom of the page.

Equation (14) can be written as

$$\sum_{g=1}^{F-1} \sum_{h>g}^{F} y_g y_h \begin{vmatrix} a_{i_1,g} & a_{i_1,h} \\ a_{i_2,g} & a_{i_2,h} \end{vmatrix} \begin{vmatrix} b_{j_1,g} & b_{j_1,h} \\ b_{j_2,g} & b_{j_2,h} \end{vmatrix} = 0. \tag{15}$$

Each quadruple (i_1, i_2, j_1, j_2), $i_1 \in \{1, \ldots, I\}$, $i_2 \in, \{1, \ldots, I\}$, $j_1 \in \{1, \ldots, J\}$, $j_2 \in \{1, \ldots, J\}$ gives

$$\sum_{g=1}^{F-1} \sum_{h>g}^{F} y_g y_h \left(a_{i_1,g} a_{i_2,h} b_{j_1,g} b_{j_2,h} + a_{i_1,h} a_{i_2,g} b_{j_1,h} b_{j_2,g} - a_{i_1,g} a_{i_2,h} b_{j_1,h} b_{j_2,g} - a_{i_1,h} a_{i_2,g} b_{j_1,g} b_{j_2,h}\right) = 0. \tag{14}$$

rise to an equation as in (15). Each such equation can be put in bilinear form[5]

$$[y_1,\ldots,y_F]\mathbf{W}_{i_1,i_2,j_1,j_2}\begin{bmatrix} y_1 \\ \vdots \\ y_F \end{bmatrix} = 0 \tag{16}$$

for $i_1 = 1,\ldots,I$, $i_2 = 1,\ldots,I$, $j_1 = 1,\ldots,J$, and $j_2 = 1,\ldots,J$.

We are now ready to state the equivalent necessary and sufficient condition on identification of restricted CP models.

Condition B: The set of equations in (16) only admits solutions satisfying $\omega([y_1,\ldots,y_F]) \leq 1$.

Note that any $\mathbf{y}$ with $\omega([y_1,\ldots,y_F]) \leq 1$ will automatically satisfy the equations in (16).

When Condition B holds, it is easily seen that $\omega(\mathbf{x}^H\mathbf{C}) \leq \omega(\mathbf{x}^H\bar{\mathbf{C}})$ for all $\omega(\mathbf{x}^H\bar{\mathbf{C}}) \leq 1$. Using Kruskal's Permutation Lemma, it follows that $\bar{\mathbf{C}}$ is the same as $\mathbf{C}$ up to permutation and column scaling, and therefore, CP decomposition is unique. When Condition B fails, there exists a vector $\mathbf{y}_0$ having weight greater than or equal to 2, such that the matrix $\mathbf{M} = \mathbf{A}\,diag(\mathbf{y}_0)\mathbf{B}^T$ is of rank at most 1. When $r_{\mathbf{M}} \leq 1$, then $\mathbf{M} = \bar{\mathbf{a}}_1\bar{\mathbf{b}}_1^T$ for certain (possibly zero) $\bar{\mathbf{a}}_1$ and $\bar{\mathbf{b}}_1$. Invoking the identity $vec(\mathbf{A}\,diag(\mathbf{x}^T)\mathbf{B}^T) = (\mathbf{A}\odot\mathbf{B})\mathbf{x}$, we conclude that a nontrivial linear combination of columns of $\mathbf{A}\odot\mathbf{B}$ can be written in the form of a tensor product of two vectors. The necessity of Condition B now follows from the earlier proof of necessity of Condition A.

Sometimes, solving a system of bilinear equations such as (16) is not as complicated as it appears. If all $\mathbf{W}_{i_1,i_2,j_1,j_2}$ are real positive semi-definite matrices, then the solutions to the system of bilinear equations can be obtained by solving a suitable linear equation. Unfortunately, this is not the case for our problem. More often than not, $\mathbf{W}_{i_1,i_2,j_1,j_2}$ are in-definite complex matrices. This poses difficulties in checking whether solutions to (16) adhere to the constraint in Condition B. We do not have a general tool for handling this verification yet, but, as will be shown shortly, some instructive simple cases can be worked out by hand, and the issue is currently under investigation.

It is also worth mentioning that instead of casting (15) into (16), we can "linearize" (15) as follows:

$$\mathbf{U}\begin{bmatrix} y_1y_2 \\ \vdots \\ y_gy_h \\ \vdots \\ y_{F-1}y_F \end{bmatrix} = 0 \tag{17}$$

where the entries of $\mathbf{U}$ are determined by (15).

In this way, we deal with a linear equation that involves a structured vector. Note that $\omega([y_1,\ldots,y_F]) \leq 1$ is *equivalent* to $[y_1y_2,\ldots,y_gy_h,\ldots,y_{F-1}y_F] = \mathbf{0}$. $\mathbf{U}$ being full column rank guarantees that $\omega([y_1,\ldots,y_F]) \leq 1$. However, a rank-deficient $\mathbf{U}$ does *not* necessarily imply that $\omega([y_1,\ldots,y_F]) \geq 2$ since $[y_1y_2,\ldots,y_{F-1}y_F]$ is a *structured* vector. In particular, simulations show that $\mathbf{U}$ can be rank-deficient even when $\mathbf{A}$ and $\mathbf{B}$ are drawn from a continuous distribution (e.g., the entries of $\mathbf{A}$ and $\mathbf{B}$ are independent and identically distributed Gaussian random variables).

[5]Note that since $\mathbf{W}_{i_1,i_2,j_1,j_2} = \mathbf{0}$ if $i_1 = i_2$ or $j_1 = j_2$, the number of active bilinear equations can be reduced.

V. Discussion

A. ten Berge's Example and Counter-Example

One of the motivations of this paper is to explain the puzzle brought by the counterexample to necessity of Kruskal's condition given in [16]. In [16], two simple examples illustrate that the uniqueness of CP decomposition depends on the particular joint pattern of zeros in the component matrices. The first example is

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & a_1 \\ 0 & 1 & 0 & a_2 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$
$$\mathbf{B} = \begin{bmatrix} 1 & 0 & 0 & b_1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & b_3 \end{bmatrix}$$

with a_1, a_2, b_1 and b_3 nonzero, and $\mathbf{C} = \mathbf{I}_4$. The second example in [16] is given by changing the first example slightly to have the zero entry in the last columns of $\mathbf{A}$ and $\mathbf{B}$ in the same place as follows:

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & a_1 \\ 0 & 1 & 0 & a_2 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$
$$\mathbf{B} = \begin{bmatrix} 1 & 0 & 0 & b_1 \\ 0 & 1 & 0 & b_2 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

with a_1, a_2, b_1, and b_2 nonzero, and a common $\mathbf{C}$. It has been proven in [16] that the decomposition of the array given by $(\mathbf{A},\mathbf{B},\mathbf{C})$ in the first example is unique, whereas alternative decompositions arise in the second example. However, no explanation on this interesting phenomenon was provided. Equipped with the main results of the previous Section, we are now in position to offer such explanation.

In the first example, we know that

$$\sum_{f=1}^{F} y_f\mathbf{a}_f\mathbf{b}_f^T = \begin{bmatrix} y_1 + a_1b_1y_4 & 0 & a_1b_3y_4 \\ a_2b_1y_4 & y_2 & a_2b_3y_4 \\ 0 & 0 & y_3 \end{bmatrix}. \tag{18}$$

Therefore, following (15), we have

$$\begin{cases} y_1y_2 + a_1b_1y_2y_4 = 0 \\ y_1y_3 + a_1b_1y_3y_4 = 0 \\ y_2y_4 = 0 \\ y_2y_3 = 0 \\ y_1y_4 = 0. \end{cases} \tag{19}$$

Equation (19) can be written as $\mathbf{U}\mathbf{w} = \mathbf{0}$, with $\mathbf{U}$ being 5×6 and $\mathbf{w} := [y_1y_2, y_1y_3, y_1y_4, y_2y_3, y_2y_4, y_3y_4]^T$. Equation (19) admits a solution of weight larger than 1 if and only if there is a nonzero $\mathbf{w}$ orthogonal to the five rows of $\mathbf{U}$. For the particular $\mathbf{U}$ in (16), the only possibility for this is to have $\mathbf{w}$ proportional to $[0,\ a_1b_1,\ 0,\ 0,\ 0, -1]$ with y_1, y_3, and y_4 nonzero. Because $\mathbf{w}_3 = y_1y_4 = 0$, this is not possible, and $\mathbf{w}$ is the zero vector after all. Therefore, (19) does not admit a solution with $\omega([y_1, y_2, y_3, y_4]) \geq 2$.

On the other hand, in the second example, we have

$$\begin{cases} y_1y_2 + a_2b_2y_1y_4 + a_1b_1y_2y_4 = 0 \\ y_1y_3 = 0 \\ y_2y_3 = 0 \\ y_3y_4 = 0. \end{cases} \tag{20}$$

It is easy to see that (20) admits the solution $y_1 = -a_1b_1/(1 + a_2b_2)$, $y_2 = 1$, $y_3 = 0$, $y_4 = 1$, where we assume $1 + a_2b_2 \neq 0$. If $1 + a_2b_2 = 0$, we can modify the values of y_2 and y_4 accordingly and still have a solution with $\omega([y_1, y_2, y_3, y_4]) \geq 2$. As it has been shown in [16], the number of such solutions is infinite.

Furthermore, in the second example

$$\mathbf{A} \odot \mathbf{B} = \begin{bmatrix} 1 & 0 & 0 & a_1b_1 \\ 0 & 0 & 0 & a_1b_2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_2b_1 \\ 0 & 1 & 0 & a_2b_2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

we can see that a linear combination of the first, the second, and the forth columns of $\mathbf{A} \odot \mathbf{B}$ constitutes a vector in the form of tensor product of two vectors as follows:

$$-\frac{a_1b_1}{1+a_2b_2}\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} a_1b_1 \\ a_1b_2 \\ 0 \\ a_2b_1 \\ a_2b_2 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{a_1a_2b_1b_2}{1+a_2b_2} \\ a_1b_2 \\ 0 \\ a_2b_1 \\ 1 + a_2b_2 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} a_2b_1 \\ 1 + a_2b_2 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} \frac{a_1b_2}{1+a_2b_2} \\ 1 \\ 0 \end{bmatrix}. \tag{21}$$

B. Bilinear Decomposition Under CM Constraints

Although bilinear decomposition is not unique in general, bilinear decomposition with CM constraints can be unique [9], [15], [17]. Interestingly, as pointed out next, the identification condition on bilinear decomposition with CM constraints is very similar to Condition B derived herein for the identification of restricted CP models. Any progress on the identification of bilinear decomposition with CM constraints might be beneficial to better understand Condition B and vice versa.

Let $\mathbf{E} = \mathbf{A}\mathbf{B}^T$, with full column rank $\mathbf{B}$ and a CM constraint on $\mathbf{A}$, that is, without loss of generality, $|a_{i,f}| = a_{i,f}a_{i,f}^H = 1$. Note that $\mathbf{E}$, $\mathbf{A}$, and $\mathbf{B}$ play the roles of $\mathbf{X}^{JI \times K}$, $(\mathbf{A} \odot \mathbf{B})$, and $\mathbf{C}$ in (6), respectively.

Like CP decomposition, bilinear decomposition with CM constraints, when unique, it is unique up to column permutation and scaling. Therefore, Kruskal's Permutation Lemma can again be taken as the cornerstone for uniqueness. Earlier work on the identification of bilinear mixtures under CM constraints [9], [15] has yielded sufficient conditions, but necessity has been left open to the best of our knowledge. Equipped with Kruskal's Permutation Lemma, we are ready to give a necessary and sufficient condition for unique bilinear decomposition under CM constraints. In this context, uniqueness means that if there is another pair $(\bar{\mathbf{A}}, \bar{\mathbf{B}})$, with $\bar{\mathbf{A}}$ having CM elements such that $\mathbf{A}\mathbf{B}^T = \bar{\mathbf{A}}\bar{\mathbf{B}}^T$, then there exists a permutation matrix $\mathbf{\Pi}$ and two nonsingular diagonal matrices $\mathbf{\Lambda}_A$, $\mathbf{\Lambda}_B$ with $\mathbf{\Lambda_A}\mathbf{\Lambda_B} = \mathbf{I}$ such that $\bar{\mathbf{A}} = \mathbf{A}\mathbf{\Pi}\mathbf{\Lambda_A}$, $\bar{\mathbf{B}} = \mathbf{B}\mathbf{\Pi}\mathbf{\Lambda_B}$. Note that the scaling indeterminacy remains despite the CM constraint, due to the possibility of rotation in the complex plane. In addition, since we have assumed a full column rank $\mathbf{B}$, it suffices to show that $\bar{\mathbf{B}} = \mathbf{B}\mathbf{\Pi}\mathbf{\Lambda_B}$ for a permutation matrix $\mathbf{\Pi}$ and a diagonal matrix $\mathbf{\Lambda}_B$; the result for $\bar{\mathbf{A}}$ then follows by simple inversion. This is the usual route taken to show uniqueness in this context.

Note that

$$\left|\sum y_f a_{i_1,f}\right| = \left|\sum y_f a_{i_2,f}\right| \Rightarrow \sum_{h=1}^{F}\sum_{g\neq h}\begin{vmatrix} a_{i_1,g} & a_{i_2,h}^H \\ a_{i_2,g} & a_{i_1,h}^H \end{vmatrix} y_g y_h^H = 0 \tag{22}$$

or

$$[y_1, \ldots, y_F]\mathbf{W}_{i_1,i_2}\begin{bmatrix} y_1^H \\ \vdots \\ y_F^H \end{bmatrix} = 0$$

$$\mathbf{W}_{i_1,i_2}(g,h) := \begin{vmatrix} a_{i_1,g} & a_{i_2,h}^H \\ a_{i_2,g} & a_{i_1,h}^H \end{vmatrix}$$

for $i_1 = 1, \ldots, I, i_2 = 1, \ldots, I$. Equation (22) follows from the CM constraints: $|a_{i,f}| = 1$. The necessary and sufficient condition for unique bilinear decomposition under CM constraints can now be stated:

Condition C: The Set of equations in (22) only admits solutions satisfying $\omega([y_1, \ldots, Y_f]) \leq 1$.

Let us show that Condition C guarantees $\bar{\mathbf{B}} = \mathbf{B}\mathbf{\Pi}\mathbf{\Lambda_B}$ for a permutation matrix $\mathbf{\Pi}$ and a diagonal matrix $\mathbf{\Lambda}_B$. Invoking the Permutation Lemma 1 , it suffices to show that $\omega(\mathbf{x}^H\mathbf{B}) \leq \omega(\mathbf{x}^H\bar{\mathbf{B}})$ for all $\omega(\mathbf{x}^H\bar{\mathbf{B}}) \leq 1$. To see this, note that Condition C guarantees $\mathbf{A}$ being full rank. This is because Condition C is equivalent to none of the nontrivial linear combinations of columns of $\mathbf{A}$ that can be written as a vector comprising CM entries, and a zero vector is a CM vector. This can be easily seen from the left side of (22). Then, from the hypothesis $\mathbf{A}\mathbf{B}^T = \bar{\mathbf{A}}\bar{\mathbf{B}}^T$ and the assumption that $\mathbf{B}$ is full column rank, it follows using Sylvester's inequality that $\bar{\mathbf{B}}$ is full column rank as well, in which case (cf. statement of Lemma 1), we only need to verify that $\omega(\mathbf{x}^H\mathbf{B}) \leq \omega(\mathbf{x}^H\bar{\mathbf{B}})$ holds for all $\omega(\mathbf{x}^H\bar{\mathbf{B}}) = 1$.

From the hypothesis $\mathbf{A}\mathbf{B}^T = \bar{\mathbf{A}}\bar{\mathbf{B}}^T$, we have

$$\bar{\mathbf{A}}\bar{\mathbf{B}}^T(\mathbf{x}^H)^T = \mathbf{A}\mathbf{B}^T(\mathbf{x}^H)^T$$

for all $\mathbf{x}$. In particular, for all those $\mathbf{x}$ such that $\omega(\mathbf{x}^H\bar{\mathbf{B}}) = 1$, the left-hand side is a column drawn from $\bar{\mathbf{A}}$ and, thus, a vector comprised of CM entries. The first element of the right-hand side is a linear combination of the elements in the first row of $\mathbf{A}$, the second is a linear combination of the elements in the second row of $\mathbf{A}$, and so on. All these row combinations should have equal modulus. If the only way for this to happen is that a single column is selected from $\mathbf{A}$, as per Condition C, then it must be that $\omega(\mathbf{x}^H\mathbf{B}) = 1$. This shows that Condition C is sufficient for uniqueness. For the converse, suppose that Condition C is violated. Without loss of generality, we may assume that $\mathbf{B}$ is an identity matrix $\mathbf{I}_F$ and that a linear combination of the first two columns of $\mathbf{A}$ constitutes a constant modulus vector $\bar{\mathbf{a}}_1$, i.e.,

$$\mathbf{a}_1 + \mathbf{a}_2 = \bar{\mathbf{a}}_1$$

and the modulus of each entry of $\bar{\mathbf{a}}_1$ is equal to a, which is a constant that is not necessarily equal to one.

If a is zero, we know $\mathbf{A}$ is rank deficient. Then, adding any nonzero null vectors of $\mathbf{A}$ to the first column of $\mathbf{B}^T$ preserves $\mathbf{A}\mathbf{B}^T$ but generates a different solution for $\mathbf{B}$.

If a is not zero, it is easy to see that

$$\begin{aligned}\mathbf{A}\mathbf{B}^T &= [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_F]\,\mathbf{I}_F \\ &= \left[\frac{\bar{\mathbf{a}}_1}{\sqrt{a}}, \mathbf{a}_2, \ldots, \mathbf{a}_F\right]\begin{bmatrix}\sqrt{a} & 0 & \mathbf{0}\\ -1 & 1 & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{F-2}\end{bmatrix} \\ &= \bar{\mathbf{A}}\bar{\mathbf{B}}^T\end{aligned} \tag{23}$$

where

$$\begin{aligned}\bar{\mathbf{A}} &= \left[\frac{\bar{\mathbf{a}}_1}{\sqrt{a}}, \mathbf{a}_2, \ldots, \mathbf{a}_F\right] \\ \bar{\mathbf{B}} &= \begin{bmatrix}\sqrt{a} & 0 & \mathbf{0}\\ -1 & 1 & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{F-2}\end{bmatrix}^T.\end{aligned} \tag{24}$$

Clearly, the modulus of each entry of $\bar{\mathbf{A}}$ is one. Hence, $(\bar{\mathbf{A}}, \bar{\mathbf{B}})$ constitutes an alternative decomposition. This completes the necessity part for Condition C.

We now can see that Condition B for the identification of restricted CP models and Condition C for the identification of bilinear models subject to CM constraints are very similar. While both have been derived using Kruskal's Permutation Lemma, they stem from conceptually very different structural constraints on the equivalent bilinear models. More specifically, CP can be viewed as a bilinear model with Khatri–Rao product structure along the row dimension, whereas CM is a bilinear model with a modulus constraint on the elements of one matrix factor.

When the CM constraint is imposed along one or more modes of CP, identifiability naturally improves in terms of the number of available equations. For instance, given a CP model $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ with full column rank $\mathbf{C}$ and CM constraints on both $\mathbf{A}$ and $\mathbf{B}$, the following is a necessary and sufficient set of uniqueness conditions:

Both

$$\sum_{g=1}^{F-1}\sum_{h>g}^{F}\begin{vmatrix}a_{i_1,g} & a_{i_1,h}\\ a_{i_2,g} & a_{i_2,h}\end{vmatrix}\begin{vmatrix}b_{j_1,g} & b_{j_1,h}\\ b_{j_2,g} & b_{j_2,h}\end{vmatrix} y_g y_h = 0$$

and

$$\sum_{h=1}^{F}\sum_{g\neq h}\begin{vmatrix}a_{i_1,g}b_{j_1,g} & a_{i_2,h}^H b_{j_2,h}^H\\ a_{i_2,g}b_{j_2,g} & a_{i_1,h}^H b_{j_1,h}^H\end{vmatrix} y_g y_h^H = 0$$

with $i_1 \in \{1,\ldots,I\}$, $i_2 \in \{1,\ldots,I\}$, $j_1 \in \{1,\ldots,J\}$, $j_2 \in \{1,\ldots,J\}$, only admit joint solutions with $\omega([y_1,\ldots,y_F]) \leq 1$.

A concise unifying treatment of necessary and sufficient uniqueness conditions for the identification of general[6] CP models subject to CM constraints along one or more modes is not available at this point. Nevertheless, individual cases can be dealt with, given the tools developed herein. Even discarding CM constraints, stating and checking necessary and sufficient uniqueness conditions for unrestricted CP models is possible but cumbersome. When none of the component matrices is full column rank and following the roadmap provided in Section III, one has to show that $\bar{\mathbf{A}} = \mathbf{A}\mathbf{\Pi}_\mathbf{A}\mathbf{\Lambda}_\mathbf{A}$ and $\bar{\mathbf{B}} = \mathbf{B}\mathbf{\Pi}_\mathbf{B}\mathbf{\Lambda}_\mathbf{B}$ separately. High-order minors of $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ must be exploited, and the condition for the identification of general CP models boils down to a number of multilinear equations with particular constraints on common solutions. We defer this pursuit at this point, pending further understanding of Condition B, which we hope to develop in on-going work.

VI. Conclusions

Two equivalent necessary and sufficient conditions for unique decomposition of restricted CP models where at least one of the component matrices is full column rank have been derived. These conditions explain the puzzle in [16]. A strong similarity between the conditions for unique decomposition of bilinear models subject to CM constraints and certain restricted CP models has been pointed out. It is hoped that this link will facilitate cross-fertilization and unification of associated uniqueness results. Last but not least, Kruskal's Permutation Lemma has been demystified. The new proof should be accessible to a much wider readership than Kruskal's original proof.

Appendix
Kruskal's Permutation Lemma: Redux

Kruskal's Permutation Lemma 1: We are given two matrices $\mathbf{A}$ and $\bar{\mathbf{A}}$, which are $I \times F$ and $I \times \bar{F}$. Suppose $\mathbf{A}$ has no zero columns. If for any vector $\mathbf{x} \in \mathbb{C}^N$ such that

$$\omega(\mathbf{x}^H\bar{\mathbf{A}}) \leq F - r_{\bar{\mathbf{A}}} + 1$$

we have

$$\omega(\mathbf{x}^H\mathbf{A}) \leq \omega(\mathbf{x}^H\bar{\mathbf{A}})$$

[6]This means without the full column rank restriction along one mode.

then $F \leq \bar{F}$; in addition, if $F \geq \bar{F}$, then $F = \bar{F}$, and there exist a permutation matrix $\mathbf{P}_{\bar{\mathbf{A}}}$ and a nonsingular matrix $\mathbf{\Lambda}$ such that $\mathbf{A} = \bar{\mathbf{A}}\mathbf{P}_{\bar{\mathbf{A}}}\mathbf{\Lambda}$.

Remark 1: Kruskal's condition is equivalent to the first equation at the bottom of the page, which implies the second equation at the bottom of the page. To show why the first statement implies the second statement, we proceed by contradiction. Suppose that there is a collection of $c_0 \geq r_{\bar{\mathbf{A}}} - 1$ columns of $\bar{\mathbf{A}}$, say, $\{\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_{c_0}\}$, and there are *only* $(c_0 - k)$ columns of $\mathbf{A}$, say, $\{\mathbf{a}_1, \ldots, \mathbf{a}_{c_0-k}\}$, such that

$$Span(\{\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_{c_0}\}) \supseteq Span(\{\mathbf{a}_1, \ldots, \mathbf{a}_{c_0-k}\}) \quad (25)$$

where $1 \leq k \leq c_0$.

Note the each of the remaining columns of $\mathbf{A}$, i.e., $\{\mathbf{a}_{c_0-k+1}, \ldots, \mathbf{a}_F\}$, is linearly independent with the column set $\{\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_{c_0}\}$; otherwise, k can be reduced by 1; this implies that for every $i = c_0 - k + 1, \ldots, F$, there exists a certain nonzero vector $\mathbf{x}_i \in Null(\{\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_{c_0}\})$ such that

$$\mathbf{x}_i^H \mathbf{a}_i \neq 0.$$

Otherwise, if for every $\mathbf{x} \in Null(\{\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_{c_0}\})$, $\mathbf{x} \in Null(\{\mathbf{a}_{i_0}\})$ for a certain $i_0 \in [c_0 - k + 1, \ldots, F]$, this implies

$$Null(\{\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_{c_0}\}) \subseteq Null(\{\mathbf{a}_{i_0}\})$$

i.e.,

$$Span(\{\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_{c_0}\}) \supseteq Span(\{\mathbf{a}_{i_0}\})$$

which means that k can be reduced by 1 as well.

Let us assume that $Null(\{\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_{c_0}\})$ is an m-dimensional linear subspace $m \geq 1$. $m = 0$ means that $Span(\{\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_{c_0}\}) = \mathbb{C}^I$; this further implies that all columns of $\mathbf{A}$ belong to $Span(\{\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_{c_0}\})$.

Now, consider

$$Null(\{\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_{c_0}, \mathbf{a}_i\})$$

for each i. Due to the existence of aforementioned $\mathbf{x}_i$, $Null(\{\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_{c_0}, \mathbf{a}_i\})$ is a proper linear subspace of $Null(\{\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_{c_0}\})$ with dimension $m-1$. Since the union of a countable number of $(m-1)$-dimensional linear subspaces of $\mathbb{C}^I$ cannot cover an m-dimensional linear subspace of $\mathbb{C}^I$, we are able to find a nonzero vector $\mathbf{x}_0 \in Null(\{\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_{c_0}\})$ such that

$$\mathbf{x}_0^H \mathbf{a}_i \neq 0, \ \forall i = c_0 - k + 1, \ldots, F.$$

The existence of such $\mathbf{x}_0$ contradicts the first statement.

Unlike the statement of Lemma 1, where the column sizes of $\bar{\mathbf{A}}$ and $\mathbf{A}$ might be different, we assume that both $\bar{\mathbf{A}}$ and $\mathbf{A}$ are $I \times F$ matrices; furthermore, without loss of generality, we assume both do not contain zero columns.

Remark 2: Given two nontrivial vectors $\bar{\mathbf{y}}$ and $\mathbf{y}$, they are linearly dependent if and only if $\omega(\mathbf{x}^H\mathbf{y}) = 0$ for all $\mathbf{x}$ satisfying $\omega(\mathbf{x}^H\bar{\mathbf{y}}) = 0$. This can be easily checked using the testing vector $\mathbf{x} = [a, 0, \ldots, 0, b, 0, \ldots, 0]$ with a, b chosen such that $\omega(\mathbf{x}^H\bar{\mathbf{y}}) = 0$.

Lemma 2: Given $\bar{\mathbf{A}} \in \mathbb{C}^{I\times F}$ and $\mathbf{A} \in \mathbb{C}^{I\times F}$, $\mathbf{A} = \bar{\mathbf{A}}\mathbf{P}_{\bar{\mathbf{A}}}\mathbf{\Lambda}$ if and only if $\omega(\mathbf{x}\mathbf{A}) \leq \omega(\mathbf{x}\bar{\mathbf{A}})$ for all $\mathbf{x}$.

Proof of Lemma 2: It suffices to prove the "if" part, and we will prove this by the induction on the number of columns of $\bar{\mathbf{A}}$, namely F.

When $F = 1$, the condition in Lemma 2 implies that $\omega(\mathbf{x}^H\mathbf{A}) = 0$ for all $\mathbf{x}$ satisfying $\omega(\mathbf{x}^H\bar{\mathbf{A}}) = 0$. From Remark 2, this implies that $\bar{\mathbf{A}}$ and $\mathbf{A}$ are linearly dependent.

Assume that Lemma 2 holds true for all $F \leq K$. Now, consider $F = K + 1$. Let $\bar{\mathbf{a}}_i$ denote the ith column of $\bar{\mathbf{A}}$, and let $\mathbf{a}_j$ denote the jth column of $\mathbf{A}$.

We claim that under the condition in Lemma 2, there must exist at least one column of $\mathbf{A}$, $\mathbf{a}_{j_0}$, which is linearly dependent with $\bar{\mathbf{a}}_1$. We will prove this by contradiction. Suppose that this claim is not true; then, based on Remark 2 and the assumption that $\mathbf{A}$ does not contain zero columns, we know that for every j, there exists a $\mathbf{x}_j$ such that

$$\omega(\mathbf{x}_j^H\bar{\mathbf{a}}_1) = 0, \ \omega(\mathbf{x}_j^H\mathbf{a}_j) = 1, \ \forall j = 1, \ldots, F. \quad (26)$$

Then, we will show that in fact there exists a *common* $\mathbf{x}_0$ such that

$$\omega(\mathbf{x}_0^H\bar{\mathbf{a}}_1) = 0, \ \omega(\mathbf{x}_0^H\mathbf{a}_j) = 1, \ \forall j = 1, \ldots, F. \quad (27)$$

$\bar{\mathbf{a}}_1 \neq \mathbf{0}$; hence, the null space of $\bar{\mathbf{a}}_1$, $Null(\bar{\mathbf{a}}_1)$ is an $(I-1)$-dimensional linear space.

Now, consider $Null(\{\bar{\mathbf{a}}_1, \mathbf{a}_j\})$ for all j. Clearly, all $Null(\{\bar{\mathbf{a}}_1, \mathbf{a}_j\})$ are covered by $Null(\bar{\mathbf{a}}_1)$. It is clearly seen that the existence of $\mathbf{x}_0$ in (27) is equivalent to $\bigcup_{j=1}^F Null(\{\bar{\mathbf{a}}_1, \mathbf{a}_j\}) \neq Null(\bar{\mathbf{a}}_1)$.

If a certain vector is orthogonal to $c \geq r_{\bar{\mathbf{A}}} - 1$ columns of $\bar{\mathbf{A}}$, then it must be orthogonal to at least c columns of $\mathbf{A}$

For every collection of $c \geq r_{\bar{\mathbf{A}}} - 1$ columns of $\bar{\mathbf{A}}$, there exists a collection of at least c columns of $\mathbf{A}$ such that

$$Span(\text{these } c \geq r_{\bar{\mathbf{A}}} - 1 \text{ columns of } \bar{\mathbf{A}}) \supseteq Span(c \text{ or more columns of } \mathbf{A}).$$

Recall that for every j, there exists a $\mathbf{x}_j$ such that

$$\omega(\mathbf{x}_j^H \bar{\mathbf{a}}_1) = 0, \;\; \omega(\mathbf{x}_j^H \mathbf{a}_j) = 1, \; \forall j$$

which implies that $Null(\{\bar{\mathbf{a}}_1, \mathbf{a}_j\})$ cannot be the same as $Null(\bar{\mathbf{a}}_1)$, but rather a proper linear subspace of $Null(\bar{\mathbf{a}}_1)$ with dimension $I-2$. Furthermore, the union of a countable number of $(I-2)$-dimensional linear subspaces of $\mathbb{C}^I$ cannot cover an $(I-1)$-dimensional subspace of $\mathbb{C}^I$; see also Fig. 1. Therefore, $\bigcup_{j=1}^{F} Null(\{\bar{\mathbf{a}}_1, \mathbf{a}_j\})$ does not cover $Null(\bar{\mathbf{a}}_1)$, and hence, we have a $\mathbf{x}_0$ such that (27) holds.

This implies that

$$\omega(\mathbf{x}_0^H \bar{\mathbf{A}}) \leq F - 1 < F = \omega(\mathbf{x}_0^H \mathbf{A})$$

which contradicts the condition in Lemma 2. Therefore, we can claim there exists at least one column of $\mathbf{A}$, which is linearly dependent with $\bar{\mathbf{a}}_1$. Without loss generality, we say this column is $\mathbf{a}_{j_0}$. Clearly

$$\omega(\mathbf{x}^H \bar{\mathbf{a}}_1) = \omega(\mathbf{x}^H \mathbf{a}_{j_0}), \forall \mathbf{x}. \tag{28}$$

Now, construct a submatrix of $\bar{\mathbf{A}}$ by removing column $\bar{\mathbf{a}}_1$ from $\bar{\mathbf{A}}$, and denote this matrix $\bar{\mathbf{A}}_0$; similarly construct a submatrix of $\mathbf{A}$ by removing column $\mathbf{a}_{j_0}$ from $\mathbf{A}$, and denote this matrix $\mathbf{A}_0$.

From $\omega(\mathbf{x}^H \mathbf{A}) \leq \omega(\mathbf{x}^H \bar{\mathbf{A}})$ for all $\mathbf{x}$ (condition in statement of Lemma 2) and (28), it follows that

$$\omega(\mathbf{x}^H \mathbf{A}_0) \leq \omega(\mathbf{x}^H \bar{\mathbf{A}}_0), \forall \mathbf{x}.$$

However, $\bar{\mathbf{A}}_0$ and $\mathbf{A}_0$ are K-column matrices; the result then follows from the induction hypothesis. That is, the $(K+1)$-column matrices $\bar{\mathbf{A}}$, $\mathbf{A}$ are the same up to permutation and scaling of columns. This completes the proof. □

Remark 3: The proof of Lemma 2 can be also applied to the following corollary.

Corollary 1: Given $\bar{\mathbf{A}} \in \mathbb{C}^{I\times F}$ and $\mathbf{A} \in \mathbb{C}^{I\times F}$, $\mathbf{A} = \bar{\mathbf{A}}\mathbf{P}_{\bar{\mathbf{A}}}\mathbf{\Lambda}$ if and only if $\omega(\mathbf{x}^H \mathbf{A}) \leq \omega(\mathbf{x}^H \bar{\mathbf{A}})$ for any $\mathbf{x}$ such that

$$\omega(\mathbf{x}^H \bar{\mathbf{A}}) \leq F - 1.$$

Compared with Kruskal's result, the conditions in both Lemma 2 and Corollary 1 appear more restrictive. There is a gap between the results presented in this Appendix so far and Kruskal's result. If $r_{\bar{\mathbf{A}}} = 2$, this gap has been filled by Corollary 1. For the general case, we have the following Lemma.

Lemma 3: Given $\bar{\mathbf{A}} \in \mathbb{C}^{I\times F}$ and $\mathbf{A} \in \mathbb{C}^{I\times F}$, $\omega(\mathbf{x}^H \mathbf{A}) \leq \omega(\mathbf{x}^H \bar{\mathbf{A}})$ for all $\mathbf{x}$ if and only if $\omega(\mathbf{x}^H \mathbf{A}) \leq \omega(\mathbf{x}^H \bar{\mathbf{A}})$ for any vector $\mathbf{x}$ such that $\omega(\mathbf{x}^H \bar{\mathbf{A}}) \leq F - r_{\bar{\mathbf{A}}} + 1$.

The interesting case occurs when $r_{\bar{\mathbf{A}}}$ is strictly less than F. Without loss of generality, we assume $r_{\bar{\mathbf{A}}} < F$.

With an additional condition, namely, $k_{\bar{\mathbf{A}}} = r_{\bar{\mathbf{A}}}$, where $k_{\bar{\mathbf{A}}}$ stands for Kruskal rank of $\bar{\mathbf{A}}$, a relatively simpler proof can be obtained as follows.

Proof of Lemma 3 – Case of $k_{\bar{\mathbf{A}}} = r_{\bar{\mathbf{A}}}$: It suffices to prove the "if" part, and we prove it by contradiction. Suppose there exists a nonzero vector $\mathbf{x}_0$, such that $\omega(\mathbf{x}_0^H \mathbf{A}) > \omega(\mathbf{x}_0^H \bar{\mathbf{A}})$, and $\omega(\mathbf{x}_0^H \bar{\mathbf{A}}) > F - r_{\bar{\mathbf{A}}} + 1$, and suppose that $\omega(\mathbf{x}_0^H \bar{\mathbf{A}})$ is the *smallest number* bigger than $F - r_{\bar{\mathbf{A}}} + 1$ in the sense that $\omega(\mathbf{x}^H \mathbf{A}) \leq \omega(\mathbf{x}^H \bar{\mathbf{A}})$ for any vector $\mathbf{x}$ such that

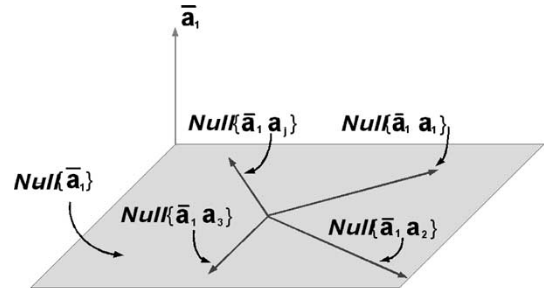


Fig. 1. Geometric illustration.

$$\omega(\mathbf{x}^H \bar{\mathbf{A}}) < \omega(\mathbf{x}_0^H \bar{\mathbf{A}}).$$

Without loss of generality, we assume

$$\omega(\mathbf{x}_0^H \bar{\mathbf{A}}) = F - r_{\bar{\mathbf{A}}} + k$$

where $2 \leq k < r_{\bar{\mathbf{A}}}$, and

$$\omega(\mathbf{x}_0^H \mathbf{A}) = F - r_{\bar{\mathbf{A}}} + k + l$$

where $1 \leq l \leq r_{\bar{\mathbf{A}}} - k$.

With such $\mathbf{x}_0$, we know that there exist $(r_{\bar{\mathbf{A}}} - k)$ columns of $\bar{\mathbf{A}}$, say, $\{\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_{r_{\bar{\mathbf{A}}}-k}\}$, and $(r_{\bar{\mathbf{A}}} - k - l)$ columns of $\mathbf{A}$, say, $\{\mathbf{a}_1, \ldots, \mathbf{a}_{r_{\bar{\mathbf{A}}}-k-l}\}$, such that

$$\{\mathbf{x}_0, \mathbf{0}\} \subseteq Null(\{\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_{r_{\bar{\mathbf{A}}}-k}\}) \bigcap Null(\{\mathbf{a}_1, \ldots, \mathbf{a}_{r_{\bar{\mathbf{A}}}-k-l}\}).$$

Since $\omega(\mathbf{x}_0^H \mathbf{A}) > \omega(\mathbf{x}_0^H \bar{\mathbf{A}})$, $\{\mathbf{a}_1, \ldots, \mathbf{a}_{r_{\bar{\mathbf{A}}}-k-l}\}$ are the only columns of $\mathbf{A}$ that can possibly belong to $Span(\{\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_{r_{\bar{\mathbf{A}}}-k}\})$; otherwise, if there is one more column, say $\mathbf{a}_F$, belonging to $Span(\{\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_{r_{\bar{\mathbf{A}}}-k}\})$, then $\mathbf{x}_0^H \mathbf{a}_F = 0$, which implies that $\omega(\mathbf{x}_0^H \mathbf{A}) = F - r_{\bar{\mathbf{A}}} + k + l - 1$ and contradicts $\omega(\mathbf{x}_0^H \mathbf{A}) = F - r_{\bar{\mathbf{A}}} + k + l$.

The remaining $F - r_{\bar{\mathbf{A}}} + k$ columns of $\bar{\mathbf{A}}$ are $\{\bar{\mathbf{a}}_{r_{\bar{\mathbf{A}}}-k+1}, \ldots, \bar{\mathbf{a}}_F\}$.

Recall that by definition of $\mathbf{x}_0$

$$\begin{aligned}\omega(\mathbf{x}^H \mathbf{A}) \leq& \,\omega(\mathbf{x}^H \bar{\mathbf{A}}) \;\; \forall \mathbf{x} \;\text{ s.t. }\; \omega(\mathbf{x}^H \bar{\mathbf{A}}) \\ \leq& \,F - r_{\bar{\mathbf{A}}} + k - 1 < \omega(\mathbf{x}_0^H \bar{\mathbf{A}}).\end{aligned} \tag{29}$$

Similar to Remark 1, we can show that (29) implies that for every $r_{\bar{\mathbf{A}}} - k + 1$ or more columns chosen from $\bar{\mathbf{A}}$, there must exist at least as many columns from $\mathbf{A}$, such that each of those from $\mathbf{A}$ is a linear combination of the said columns of $\bar{\mathbf{A}}$.

Now, consider the following $F - r_{\bar{\mathbf{A}}} + k$ column sets drawn from $\bar{\mathbf{A}}$

$$\{\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_{r_{\bar{\mathbf{A}}}-k}, \bar{\mathbf{a}}_i\}$$

where $i = r_{\bar{\mathbf{A}}} - k + 1, \ldots, F$. Each of them has $r_{\bar{\mathbf{A}}} - k + 1$ distinct columns of $\bar{\mathbf{A}}$.

According to (29), for each column set $\{\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_{r_{\bar{\mathbf{A}}}-k}, \bar{\mathbf{a}}_i\}$, there exist at least $r_{\bar{\mathbf{A}}} - k + 1$ columns $\{\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_{r_{\bar{\mathbf{A}}}-k+1}}\}$ such that each column from the latter set is a linear combination of those in the former set.

Recall that except for $\{\mathbf{a}_1, \ldots, \mathbf{a}_{r_{\bar{\mathbf{A}}}-k-l}\}$, there is no other column of $\mathbf{A}$, which belongs to $Span(\{\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_{r_{\bar{\mathbf{A}}}-k}\})$. This implies that *at least* $(l+1) = (r_{\bar{\mathbf{A}}} - k + 1) - (r_{\bar{\mathbf{A}}} - k - l)$ columns from $\{\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_{r_{\bar{\mathbf{A}}}-k+1}}\}$, other than those in $\{\mathbf{a}_1, \ldots, \mathbf{a}_{r_{\bar{\mathbf{A}}}-k-l}\}$, must be such that each is a linear combination of $\bar{\mathbf{a}}_i$ and some or all of $\{\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_{r_{\bar{\mathbf{A}}}-k}\}$. Let ϕ_i denote the column set consisting of those $l+1$ columns from $\{\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_{r_{\bar{\mathbf{A}}}-k+1}}\}$.

We claim that every two ϕ_i and ϕ_j are disjoint for $i \neq j$; if there exists a common element between ϕ_i and ϕ_j, say $\mathbf{a}_j^i \in \phi_i \bigcap \phi_j$, then

$$\mathbf{a}_j^i = \sum_{n=1}^{r_{\bar{\mathbf{A}}}-k} c_n \bar{\mathbf{a}}_n + c_i \bar{\mathbf{a}}_i = \sum_{m=1}^{r_{\bar{\mathbf{A}}}-k} d_m \bar{\mathbf{a}}_m + d_j \bar{\mathbf{a}}_j, \; c_i \neq 0, \; d_j \neq 0$$

which in turn implies that the $r_{\bar{\mathbf{A}}} - k + 2$ column set

$$\{\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_{r_{\bar{\mathbf{A}}}-k}, \bar{\mathbf{a}}_i, \bar{\mathbf{a}}_j\}$$

is a linearly dependent set of columns with distinct indices. Since $k \geq 2$ and $k_{\bar{\mathbf{A}}} = r_{\bar{\mathbf{A}}}$, we have a contradiction. Therefore, every two ϕ_i and ϕ_j are disjoint for $i \neq j$. In addition, it is easily seen that ϕ_i is disjoint with $\{\mathbf{a}_1, \ldots, \mathbf{a}_{r_{\bar{\mathbf{A}}}-k-l}\}$ as well.

The remainder is a counting problem. The number of all columns of $\mathbf{A}$ should not be less than the number of columns in all the above disjoint column subsets of $\mathbf{A}$. However, from $\{\mathbf{a}_1, \ldots, \mathbf{a}_{r_{\bar{\mathbf{A}}}-k-l}\}$, we have $r_{\bar{\mathbf{A}}} - k - l$ columns, from each ϕ_i, we have at least $l+1$ columns, and we have $F - r_{\bar{\mathbf{A}}} + k$ such ϕ_i; therefore, the total number of columns from all disjoint column subsets of $\mathbf{A}$ is not less then

$$r_{\bar{\mathbf{A}}} - k - l + (l+1)(F - r_{\bar{\mathbf{A}}} + k) = l(F - r_{\bar{\mathbf{A}}}) + F + (k-1)l$$

which is strictly greater than F for $l \geq 1$, and $k \geq 2$, whereas $\mathbf{A}$ has F columns only. We have a contradiction. □

The above proof of the special case of Lemma 3 provides helpful intuition. Armed with this insight, the following proof of Lemma 3 becomes natural.

Proof of Lemma 3—General Case: The spirit of the proof follows the earlier argument for the special case. In particular, we argue by contradiction.

Suppose that there exists a $\mathbf{x}_0$, such that $\omega(\mathbf{x}_0^H \mathbf{A}) > \omega(\mathbf{x}_0^H \bar{\mathbf{A}})$, and $\omega(\mathbf{x}_0^H \bar{\mathbf{A}}) > F - r_{\bar{\mathbf{A}}} + 1$, i.e.,

$$\omega(\mathbf{x}_0^H \bar{\mathbf{A}}) = F - r_{\bar{\mathbf{A}}} + k > F - r_{\bar{\mathbf{A}}} + 1, \;\; 2 \leq k < r_{\bar{\mathbf{A}}}$$
$$\omega(\mathbf{x}_0^H \mathbf{A}) = F - r_{\bar{\mathbf{A}}} + k + l, \;\; 1 \leq l \leq r_{\bar{\mathbf{A}}} - k$$

and suppose $\omega(\mathbf{x}_0^H \bar{\mathbf{A}}) = F - r_{\bar{\mathbf{A}}} + k$ is the *smallest number* bigger than $F - r_{\bar{\mathbf{A}}} + 1$ in the sense that $\omega(\mathbf{x}^H \mathbf{A}) \leq \omega(\mathbf{x}^H \bar{\mathbf{A}})$ for any vector $\mathbf{x}$ such that

$$\omega(\mathbf{x}^H \bar{\mathbf{A}}) \leq F - r_{\bar{\mathbf{A}}} + k - 1 < \omega(\mathbf{x}_0^H \bar{\mathbf{A}})$$

which implies that for every $r_{\bar{\mathbf{A}}} - k + 1$ or more columns chosen from $\bar{\mathbf{A}}$, there must exist at least as many columns from $\mathbf{A}$, such that each of those from $\mathbf{A}$ is a linear combination of the said columns of $\bar{\mathbf{A}}$.

As before, with such $\mathbf{x}_0$, we know that there exist $(r_{\bar{\mathbf{A}}} - k)$ columns of $\bar{\mathbf{A}}$, say, $\{\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_{r_{\bar{\mathbf{A}}}-k}\}$, and $(r_{\bar{\mathbf{A}}} - k - l)$ columns of $\mathbf{A}$, say, $\{\mathbf{a}_1, \ldots, \mathbf{a}_{r_{\bar{\mathbf{A}}}-k-l}\}$, such that

$$\{\mathbf{x}_0, \mathbf{0}\} \subseteq Null(\{\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_{r_{\bar{\mathbf{A}}}-k}\}) \bigcap Null(\{\mathbf{a}_1, \ldots, \mathbf{a}_{r_{\bar{\mathbf{A}}}-k-l}\})$$

and since $\omega(\mathbf{x}_0^H \mathbf{A}) > \omega(\mathbf{x}_0^H \bar{\mathbf{A}})$, $\{\mathbf{a}_1, \ldots, \mathbf{a}_{r_{\bar{\mathbf{A}}}-k-l}\}$ are the only columns of $\mathbf{A}$ that can possibly belong to $Span(\{\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_{r_{\bar{\mathbf{A}}}-k}\})$.

What we are going to do next is different from the previous proof. We are going to partition the remaining $F - r_{\bar{\mathbf{A}}} + k$ columns of $\bar{\mathbf{A}}$, namely, $\{\bar{\mathbf{a}}_{r_{\bar{\mathbf{A}}}-k+1}, \ldots, \bar{\mathbf{a}}_F\}$. Notice that none of those remaining columns of $\bar{\mathbf{A}}$ is going to be linearly dependent with $\{\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_{r_{\bar{\mathbf{A}}}-k}\}$; otherwise, k can be reduced by 1. We will partition those remaining $F - r_{\bar{\mathbf{A}}} + k$ columns into[7] $M \geq 2$ nonempty disjoint subsets in the sense that each subset contains one particular remaining column and all the other columns that are the linear combinations of this particular remaining column and $\{\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_{r_{\bar{\mathbf{A}}}-k}\}$. Let $S_i \geq 1$ denote the number of columns in the ith partition set. Clearly

$$\sum_{i=1}^{M} S_i = F - r_{\bar{\mathbf{A}}} + k.$$

Now, add each partition set to $\{\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_{r_{\bar{\mathbf{A}}}-k}\}$ to form a concatenated set. Each concatenated column set of $\bar{\mathbf{A}}$ has $S_i + r_{\bar{\mathbf{A}}} - k \geq r_{\bar{\mathbf{A}}} - k + 1$ columns. Recall that for every $r_{\bar{\mathbf{A}}} - k + 1$ or more columns chosen from $\bar{\mathbf{A}}$, there must exist at least as many columns from $\mathbf{A}$, such that each of those from $\mathbf{A}$ is a linear combination of the said columns of $\bar{\mathbf{A}}$. Then, there must exist at least $(S_i + r_{\bar{\mathbf{A}}} - k)$ columns of $\mathbf{A}$ such that each of those from $\mathbf{A}$ is a linear combination of those columns of the concatenated set. We already know that $\{\mathbf{a}_1, \ldots, \mathbf{a}_{r_{\bar{\mathbf{A}}}-k-l}\}$ are the only columns of $\mathbf{A}$ that can possibly belong to $Span(\{\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_{r_{\bar{\mathbf{A}}}-k}\})$; therefore, every such $(S_i + r_{\bar{\mathbf{A}}} - k)$-column subset of $\mathbf{A}$ must have at least $S_i + l = (S_i + r_{\bar{\mathbf{A}}} - k) - (r_{\bar{\mathbf{A}}} - k - l)$ columns, other than those in $\{\mathbf{a}_1, \ldots, \mathbf{a}_{r_{\bar{\mathbf{A}}}-k-l}\}$, such that each column is a linear combination of at least one column from the ith partition set and $\{\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_{r_{\bar{\mathbf{A}}}-k}\}$.

Let ϕ_i denote the column set consisting of those $S_i + l$ columns of $\mathbf{A}$.

We claim that all ϕ_i and $\{\mathbf{a}_1, \ldots, \mathbf{a}_{r_{\bar{\mathbf{A}}}-k-l}\}$ are mutually disjoint.

Suppose this is not true. Recall that no element of ϕ_i belongs to $\{\mathbf{a}_1, \ldots, \mathbf{a}_{r_{\bar{\mathbf{A}}}-k-l}\}$; hence, there is no common element between $\{\mathbf{a}_1, \ldots, \mathbf{a}_{r_{\bar{\mathbf{A}}}-k-l}\}$ and any particular ϕ_{i0}; meanwhile, if there is a column belonging to two different ϕ_i, this will contradict the way we partition the remaining columns of $\bar{\mathbf{A}}$.

[7] M can be equal to 1 only if $r_{\bar{\mathbf{A}}} = 2$; however, the case $r_{\bar{\mathbf{A}}} = 2$ has been solved by Corollary 1. For $r_{\bar{\mathbf{A}}} \geq 3$, M cannot be 1. Suppose $M = 1$; then, according to the definition of our partition, each remaining column of $\bar{\mathbf{A}}$ is the linear combination of $\bar{\mathbf{a}}_{r_{\bar{\mathbf{A}}}-k+1}$ and $\{\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_{r_{\bar{\mathbf{A}}}-k}\}$. Therefore, adding all remaining columns of $\bar{\mathbf{A}}$ to $\{\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_{r_{\bar{\mathbf{A}}}-k}\}$ could only increase the rank of $[\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_{r_{\bar{\mathbf{A}}}-k}]$ by *at most* one. This implies the rank of $\bar{\mathbf{A}}$ is bounded by $r_{\bar{\mathbf{A}}} - k + 1$, which is less than $r_{\bar{\mathbf{A}}}$ since $k \geq 2$. Hence, $M \geq 2$.

The remainder is again a counting problem. Each ϕ_i contributes at least $S_i + l$ columns. $\{\mathbf{a}_1, \ldots, \mathbf{a}_{r_{\bar{\mathbf{A}}}-k-l}\}$ also contributes $r_{\bar{\mathbf{A}}} - k - l$ columns. Summing up, we know $\mathbf{A}$ should have *at least*

$$r_{\bar{\mathbf{A}}} - k - l + \sum_{i=1}^{M}(S_i + l) = F + (M-1)l$$

columns. Since $M \geq 2$ and $l \geq 1$

$$F + (M-1)l > F$$

whereas $\mathbf{A}$ only has F columns. Hence, we have a contradiction. □

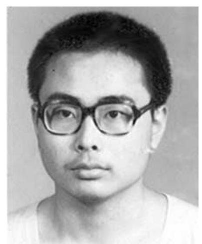
One natural question that arises at this point is whether one can further improve Lemma 1 in the sense that Lemma 1 can be viewed as an improved version of Lemma 2. Does the conclusion of Lemma 1 hold if we pose a smaller bound on the right-hand side of (7)? The answer is *no* in general. It is known that $k_{\mathbf{A}} = r_{\mathbf{A}}$ almost surely when $\mathbf{A}$ is drawn from a continuous distribution. With the aid of Remark 1, it can be seen that given a matrix $\bar{\mathbf{A}}$ with $k_{\bar{\mathbf{A}}} = r_{\bar{\mathbf{A}}}$, even if $\omega(\mathbf{x}^H\mathbf{A}) \leq \omega(\mathbf{x}^H\bar{\mathbf{A}})$ for any vectors with $\omega(\mathbf{x}^H\bar{\mathbf{A}}) \leq F - r_{\bar{\mathbf{A}}}$, $\mathbf{A}$ and $\bar{\mathbf{A}}$ are *not* necessarily equivalent up to permutation and scaling. The Lemma can be relaxed when $k_{\bar{\mathbf{A}}} = 1$, but this is not the case of interest.

ACKNOWLEDGMENT

The authors would like to thank J. M. F. ten Berge for his critical reading and helpful comments on an earlier draft of this paper.

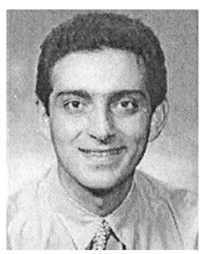
REFERENCES

[1] J. W. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Trans. Circuits Syst.*, vol. CAS-25, no. 9, pp. 772–781, Sept. 1978.
[2] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra* Philadelphia, PA, 2000.
[3] J. D. Carrol and J. J. Chang, "Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckardt-Young" decomposition," *Psychometrika*, vol. 35, pp. 283–319, 1970.
[4] R. Harshman, "Foundations of PARAFAC procedure: Models and conditions for an 'Explanatory' multi-mode factor analysis," *UCLA Working Papers Phonetics*, vol. 16, pp. 1–84, 1970.
[5] W. P. Krjnen, *The Analysis of Three-Way Arrays by Constrained PARAFAC Methods*. Leiden, Germany: DSWO, 1993.
[6] T. Jiang and N. D. Sidropoulos, "Blind identification of out-of-cell users in DS-CDMA: An algebraic approach," in *Proc. ICASSP*, Orlando, FL, May 13–17, 2002.
[7] ——, "A direct semi-blind receiver for SIMO and MIMO OFDM systems subject to frequency offset," in *Proc. SPAWC*, Rome, Italy, June 15 – 18, 2003.
[8] J. B. Kruskal, "Three-way arrays: Rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics," *Linear Algebra Applications*, vol. 18, pp. 95–138, 1977.
[9] A. Leshem, N. Petrochilos, and A. van der Veen, "Finite sample identifiability of multiple constant modulus sources," in *Proc. IEEE Sensor Array Multichannel Signal Process. Workshop*, 2002, pp. 408–412.
[10] X. Liu and N. D. Sidiropoulos, "Cramer-Rao lower bounds for low-rank decomposition of multidimensional arrays," *IEEE Trans. Signal Processing*, vol. 49, pp. 2074–2086, Sept. 2001.
[11] X. Liu, N. D. Sidiropoulos, and A. Swami, "High resolution localization and tracking of multiple frequency hopped signals," *IEEE Trans. Signal Processing*, vol. 50, pp. 891–901, Apr. 2002.
[12] N. D. Sidiropoulos and R. Bro, "On the uniqueness of multilinear decomposition of N-way arrays," *J. Chemometrics, Special Cross-Disciplinary Issue Multi-Way Anal.*, vol. 14, no. 3, pp. 229–239, May 2000.
[13] N. D. Sidiropoulos, R. Bro, and G. B. Giannakis, "Parallel factor analysis in sensor array processing," *IEEE Trans. Signal Processing*, vol. 48, pp. 2377–2388, Aug. 2000.
[14] N. D. Sidiropoulos, G. B. Giannakis, and R. Bro, "Blind PARAFAC receivers for DS-CDMA systems," *IEEE Trans. Signal Processing*, vol. 48, pp. 810–823, Mar. 2000.
[15] S. Talwar, M. Viberg, and A. Paulraj, "Blind separation of synchronous co-channel digital signals using an antenna array – Part I: Algorithms," *IEEE Trans. Signal Processing.*, vol. 44, pp. 1184–1196, May 1996.
[16] J. M. F. ten Berge and N. D. Sidiropoulos, "On uniqueness in CANDECOMP/PARAFAC," *Psychometrika*, vol. 67, 2002.
[17] A. van der Veen, "Asymptotic properties of the algebraic constant modulus algorithm," *IEEE Trans. Signal Processing.*, vol. 49, pp. 1796–1807, Aug. 2001.



Tao Jiang (S'01) received the B.S. degree from Peking University, Beijing, China, in 1997 and the M.S. degree from the University of Minnesota, Minneapolis, in 2000, both in mathematics. He is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, University of Minnesota.

His research interests are in the area of signal processing for communications, focusing on wireless communications.



Nicholas D. Sidiropoulos (M'92–SM'99) received the Diploma in electrical engineering from the Aristotelian University of Thessaloniki, Thessaloniki, Greece, and the M.S. and Ph.D. degrees in electrical engineering from the University of Maryland, College Park (UMCP), in 1988, 1990 and 1992, respectively.

From 1988 to 1992, he was a Fulbright Fellow and a Research Assistant at the Institute for Systems Research (ISR), UMCP. From September 1992 to June 1994, he served his military service as a Lecturer in the Hellenic Air Force Academy. From October 1993 to June 1994, he also was a member of the technical staff, Systems Integration Division, G-Systems Ltd., Athens, Greece. He was a Postdoctoral Fellow (1994 to 1995) and Research Scientist (1996 to 1997) at ISR-UMCP, an Assistant Professor with the Department of Electrical Engineering, University of Virginia, Charlottesville, from 1997 to 1999, and an Associate Professor with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, from 2000 to 2002. He is currently a Professor with the Telecommunications Division of the Department of Electronic and Computer Engineering, Technical University of Crete, Chania, Crete, Greece, and Adjunct Professor at the University of Minnesota. His current research interests are primarily in signal processing for communications, and multi-way analysis. He is an active consultant for industry in the areas of frequency hopping systems and signal processing for xDSL modems.

Dr. Sidiropoulos is a member of the Signal Processing for Communications Technical Committee (SPCOM-TC) of the IEEE SP Society and currently serves as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING. From 2000 to 2002, he also served as Associate Editor for the IEEE SIGNAL PROCESSING LETTERS. He received the NSF/CAREER award (Signal Processing Systems Program) in June 1998 and an IEEE Signal Processing Society Best Paper Award in 2001.