A CANDECOMP/PARAFAC Perspective on Uniqueness of DOA Estimation Using a Vector Sensor Array

Xijing Guo, Sebastian Miron, David Brie, Shihua Zhu and Xuewen Liao

Abstract

We address the uniqueness problem in estimating the directions-of-arrival (DOA's) of multiple narrowband and fully polarized signals impinging on a passive sensor array composed of identical vector sensors. The data recorded on such an array present the so-called “multiple invariances”, which can be linked to the CANDECOMP/PARAFAC (CP) model. CP refers to a family of low-rank decompositions of three-way or higher way (multidimensional) data arrays, where each dimension is termed as a “mode”. A sufficient condition is derived for uniqueness of the CP decomposition of a three-way (three mode) array in the particular case where one of the three loading matrices, each associated to one mode, involved in the decomposition has full column rank. Based on this, upper bounds on the maximal number of identifiable DOA's are deduced for the two typical cases, i.e., the general case of uncorrelated or partially correlated sources and the case where the sources are coherent.

Index Terms

Vector sensor array processing, identifiability, CANDECOMP/PARAFAC uniqueness, polarization

I. INTRODUCTION

The notion “vector sensor” was formally introduced by Nehorai and Paldi in [1], and, by contrast, the conventional ones are commonly called “scalar sensors”. A typical “complete” electromagnetic (EM) vector sensor consists of

The work of X. Guo, S. Zhu, and X. Liao was supported by the National Natural Science Foundation of China under Grant 60902043 and the National HighTechnology Research and Development Program (“863” Program) of China under Grant 2009AA011502. Part of the work (Theorem 1 and the proof) was presented at the 22nd GRETSI Symposium on Signal and Image Processing, Dijon, France, September 8-11, 2009. Copyright (c) 2011 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

X. Guo is with the Department of Information and Communication Engineering, Xi’an Jiaotong University, 710049 Xi’an, China, and also with Centre de Recherche en Automatique de Nancy (CRAN), Nancy-Université, CNRS, Boulevard des Aiguilletes B.P. 239 F-54506 Vandœuvre-lès-Nancy, France (email:xijing.guo@cran.uhp-nancy.fr). He performed the work while at the CRAN.

S. Miron and D. Brie are with Centre de Recherche en Automatique de Nancy (CRAN), Nancy-Université, CNRS, Boulevard des Aiguilletes B.P. 239 F-54506 Vandœuvre-lès-Nancy, France (email: sebastian.miron@cran.uhp-nancy.fr; david.brie@cran.uhp-nancy.fr).

S. Zhu and X. Liao are with the Department of Information and Communication Engineering, Xi’an Jiaotong University, 710049 Xi’an, China.
two orthogonal triads of dipole and loop antennas with the same phase center, yielding a vector output containing the measurements of all the six components of electromagnetic field incident on the sensor. Therefore, an EM vector sensor is intrinsically a polarization-diverse scalar sensor array, whereas it is different from those conventional arrays, such as those studied in [2] and [3], because, for a vector sensor, all the elements are configured to share a common phase center. The EM vector sensors enable resolving the DOA’s of incident signals with arbitrary polarizations, a challenging problem for the conventional scalar sensor arrays with all identically polarized sensors. They gained popularity for improving the performance of those high-resolution eigenstructure-based techniques, such as ESPRIT [4]-[9], MUSIC [10]-[12] and its recent variants [13]-[16], in disambiguating the DOA’s of the superimposed signals.

In this context, a problem of primary importance is to find the conditions under which the DOA’s can be uniquely localized. It was shown [17]-[19] that uniqueness is closely related to linear dependence of the steering vectors. Some results on linear dependence of the steering vectors for the conventional sensor arrays can be found in [20]-[23]. In particular, a detailed characterization of linear dependence of the steering vectors for a single complete vector sensor is given in [24] while in [25] the case of dipole triad is investigated. However, the problem becomes more complicated when an array comprising multiple vector sensors is considered and only few studies exist on linear dependence of the steering vectors of a vector sensor array. Among them, [26] provides a link between linear dependence of the steering vectors of a vector sensor array and that of a scalar sensor array having the same sensor configuration. These results are extended in a recent study [27] for analyzing a virtual array manifold when higher-order statistics are used with EM vector sensors.

This paper proposes an original approach for studying DOA identifiability with vector sensor arrays based on a CANDECOMP/PARAFAC (CP) model of the data. The idea of analyzing data obtained from a sensor array with multiple invariance properties (see [28] for multiple invariance) by CP was first introduced by Sidiropoulos et al. [29]. The CP decomposition, introduced independently by Carroll and Chang [30] and Harshman [31] has been largely used lately in various domains because of its attractive identifiability properties under some mild conditions. For a general overview of CP and its applications see [32] and the references therein. Based on Kruskal’s condition [33], Sidiropoulos et al. [29] also investigated identifiability of the data model for the conventional sensor array systems characterized by multiple invariance. Observing that a sensor array with all identical vector sensors also possesses multiple invariance, the results of [29] can be extended to the case of vector sensor arrays [34].

Herein we present a sufficient condition (Theorem 1) on uniqueness of the CP decomposition with a full column rank loading matrix. Some of these results have been presented previously in [35]. The same uniqueness condition was also provided independently by Stegeman in a recent work [36]. However, the proof presented in this paper is, to some extent, more condensed than that of [36]. It makes use of a new result (Lemma 1) on the steering vectors of a vector sensor array. This easy-to-check condition allows to study DOA identifiability for polarized sources measured on a vector sensor array in two typical scenarios. The first one addresses the case of uncorrelated or partially correlated sources, which is given by a full rank signal matrix. Coupled with uniqueness of CP, an upper bound on the number of resolvable sources is derived. The second scenario considers the case of coherent signals...
for which we propose a CP formulation of the polarization smoothing algorithm (PSA) of [37]. Then we derive a sufficient condition ensuring uniqueness of the DOA localization.

The remainder of the paper is organized as follows. In Section II we introduce the vector sensor model and establish the link to CP. In Section III, we briefly state some important existing results on identifiability of the CP model, and then we present the main results of the paper. Section IV applies the identifiability results to the two DOA estimation scenarios mentioned above. Finally, conclusions are drawn in Section V.

II. SYSTEM MODEL

Let us consider a sensor array system of \( M \) displaced but otherwise identical EM vector sensors deployed in the far-field of \( K \) narrowband sources. Fig. 1 illustrates one of these vector sensors. In accordance, a Cartesian coordinate system is established where the reference vector sensor is positioned at the origin. It is assumed that the signals are completely polarized and the propagation medium is isotropic and homogeneous. Each of these vector sensors forms a subarray of the entire sensor system, with the manifold given by [1]

\[
\Theta(\phi, \psi) = \begin{bmatrix}
-\sin \phi & -\cos \phi \sin \psi \\
\cos \phi & -\sin \phi \sin \psi \\
0 & \cos \psi \\
-\cos \phi \sin \psi & \sin \phi \\
-\sin \phi \sin \psi & -\cos \phi \\
\cos \psi & 0
\end{bmatrix}
\]

(1)

\[
b(\phi, \psi, \alpha, \beta) \triangleq 
\begin{bmatrix}
\cos \alpha & \sin \alpha \\
-\sin \alpha & \cos \alpha \\
0 & j \sin \beta
\end{bmatrix}
\]

\[
p(\alpha, \beta)
\]

for an incident signal with DOA \((\phi, \psi)\) and polarization \((\alpha, \beta)\). In (1), \( \phi \in [0, 2\pi) \) denotes the azimuth angle, \( \psi \in [-\pi/2, \pi/2] \) is the elevation angle, \( \alpha \in (-\pi/2, \pi/2] \) is the orientation angle of the polarization ellipse and \( \beta \in [-\pi/4, \pi/4] \) is the ellipticity angle. Throughout the paper, we assume that the DOA’s and the polarizations are unknown but deterministic.

We denote by a set of three-dimensional vectors \( \{r_m\}_{m=1}^M \), the \( M \) observation points where the vector sensors are positioned in the reference frame, and by \( u(\phi, \psi) = [\cos \phi \cos \psi \sin \phi \cos \psi \sin \psi]^T \) the unit Poynting vector in the source direction. The expression for the phase shifts induced by the displacements of the \( m \)-th \((m = 1, \ldots, M)\) vector sensor relative to the reference one is then given by \( a_m(\phi, \psi) = \exp\{j 2\pi r_m^T u(\phi, \psi)/\lambda\} \), where \( \lambda \) is the wavelength. Furthermore, define \( a(\phi, \psi) = [a_1(\phi, \psi), \ldots, a_M(\phi, \psi)]^T \), the steering vector of a virtual scalar sensor array that has the same sensor configuration as the vector sensor array [26]. Correspondingly, the steering vector of the entire vector sensor array system is given by [1] \( d(\phi, \psi, \alpha, \beta) = a(\phi, \psi) \otimes b(\phi, \psi, \alpha, \beta) \) where \( \otimes \) denotes the Kronecker product. Now suppose that the \( k \)-th signal impinges on the array from a direction \((\phi_k, \psi_k)\), with the polarization parameterized by \((\alpha_k, \beta_k)\). Inductively, the output of the array at time \( t \), i.e., \( y(t) \), is obtained by
Fig. 1. Geometry of a vector sensor: \( \{e_x, e_y, e_z\} \) denotes the dipole triad and \( \{h_x, h_y, h_z\} \) the loop triad.

summing up the contributions from all the \( K \) incident signals \( s_k(t) \):

\[
y(t) = \sum_{k=1}^{K} s_k(t) d(\phi_k, \psi_k, \alpha_k, \beta_k) = Ds(t) \tag{2}
\]

where \( D = [d(\phi_1, \psi_1, \alpha_1, \beta_1), \ldots, d(\phi_K, \psi_K, \alpha_K, \beta_K)] \) is a \( 6M \times K \) matrix, and \( s(t) = [s_1(t), \ldots, s_K(t)]^T \) is a \( K \)-element vector. As we are investigating identifiability of the model (2), only the noise free system is considered since the uniqueness problem is decoupled from the estimation problem by its nature [17]. Let \( A \triangleq [a(\phi_1, \psi_1), \ldots, a(\phi_K, \psi_K)] \) and \( B \triangleq [b(\phi_1, \psi_1, \alpha_1, \beta_1), \ldots, b(\phi_K, \psi_K, \alpha_K, \beta_K)] \). Then, one observes \( D = A \odot B \) where \( \odot \) denotes the Khatri-Rao (column-wise Kronecker) product. In the sequel, we will occasionally omit the explicit dependence on \( (\phi_k, \psi_k, \alpha_k, \beta_k) \), for notational simplicity. If \( N \) snapshots at the discrete time instants \( \{t_n\}_{n=1}^{N} \) are collected, the available data can be organized into a \( 6M \times N \) matrix \( Y = [y(t_1), \ldots, y(t_N)] = DS^T \), where \( S \triangleq [s(t_1), \ldots, s(t_N)]^T \). Observing the multiple invariances of the data model, \( Y \) can also be expressed as

\[
Y \triangleq \begin{bmatrix} Y_1 \\ \vdots \\ Y_M \end{bmatrix} = \begin{bmatrix} BD_1(A) \\ \vdots \\ BD_M(A) \end{bmatrix} S^T = (A \odot B)S^T \tag{3}
\]

where \( D_m(A) = \text{diag}(a_{m1}, \ldots, a_{mK}) \) denotes the diagonal matrix with the \( m \)th row of \( A = [a_{mk}]_{M \times K} \) as its diagonal, and \( Y_m = BD_m(A)S^T \). The factorization model (3) explicitly expresses an “unfolded” version of the three-way CP model [29], [31] for the data obtained with a vector sensor array.

III. Uniqueness Issues for the Three-way CP Decomposition

In this section, we start by introducing some basic concepts on the CP decomposition. A three-way array (or tensor) \( X \) of size \( M \times P \times N \) with typical element \( x_{mpn} \) is termed as a “rank-1” three-way array if it is given by the “outer product” of the three vectors \( a, b \) and \( c \) with the typical elements \( a_m, b_p \) and \( c_n \), respectively, in the element-wise expression \( x_{mpn} = a_m b_p c_n \). Moreover, the three-way array \( X \) is rank-\( K \) if \( K \) is the minimum
number of rank-1 tensors in the CP decomposition

\[ x_{mnp} = \sum_{k=1}^{K} a_{mk} b_{pk} c_{nk} \]  

(4)

where \( a_{mk}, b_{pk}, c_{nk} \) are the typical elements of the three “loading matrices” \( A (M \times K) \), \( B (P \times K) \) and \( C (N \times K) \) for each mode. The factorization (4) can also be expressed in the form of “slices” along the first mode as \( X_m = BD_m(A)C^T, m = 1, \ldots, M \) and in the unfolded matrix form (cf.(3))

\[
X \triangleq \begin{bmatrix}
X_1 \\
\vdots \\
X_M
\end{bmatrix} = (A \odot B)C^T. 
\]  

(5)

The CP decomposition of a tensor presents interesting uniqueness\(^1\) properties making it attractive in a wide range of applications. A most well-known uniqueness condition is due to Kruskal [33], relying on the concept of “Kruskal-rank”, or simply, \( k \)-rank.

The \( k \)-rank of a matrix \( A \in \mathbb{C}^{M \times K} \) equals \( k_A \), if \( k_A \) is the maximum number of \( \ell \) such that every \( \ell \) columns of \( A \) are linearly independent.

The link between the rank of a matrix \( A \), denoted by \( \text{rank}(A) \), and its \( k \)-rank is given by the inequality \( k_A \leq \text{rank}(A) \leq \min(M, K) \); the equalities hold if \( A \) has full rank.

Kruskal’s condition [33]

\[ k_A + k_B + k_C \geq 2K + 2 \]  

(6)

provides a sufficient condition for uniqueness of the CP decomposition (5). Moreover, it becomes a necessary and sufficient condition for \( K = 2, 3 \) [38]. For the case when one of the loading matrices, say \( C \), is full column rank, Jiang and Sidiropoulos [39] derived a necessary and sufficient condition for identifiability of the CP model, which can be stated as

The CP decomposition is unique if none of the nontrivial\(^2\) linear combinations of the columns of \( A \odot B \) can be written as a tensor product of two vectors\(^3\).

An easier-to-check, sufficient condition for uniqueness of the CP model with a full column rank loading matrix \( C \) was also provided in [39] (see also [36]). It depends on a prespecified \( M^2P^2 \times K(K-1)/2 \) matrix \( U \) whose elements are determined by the product of the second-order minors of \( A \) and \( B \) (see [39] for the details on how to generate \( U \)). Jiang and Sidiropoulos [39] have proven that if \( U \) has full column rank then identifiability of the model is guaranteed. The same condition was derived independently by De Lathauwer in [40].

\(^1\)Here, by uniqueness, we mean “essential uniqueness”, that is, if another set of matrices \( (\bar{A}, \bar{B}, \bar{C}) \) also satisfy (5) exactly, there exists a permutation matrix \( \Pi \) and three invertible diagonal scaling matrices \( \Delta_1, \Delta_2, \Delta_3 \) satisfying \( \Delta_1 \Delta_2 \Delta_3 = I \) where \( I \) is the identity matrix, such that \( \bar{A} = A\Pi\Delta_1, \bar{B} = B\Pi\Delta_2, \bar{C} = C\Pi\Delta_3. \)

\(^2\)The nontrivial linear combination is referred to as a combination involving at least two vectors.

\(^3\)The tensor product of two vectors \( a \) and \( b \) is given by their element-wise product (Kronecker product) \( a \otimes b \).
The necessary and sufficient condition of [39] is the cornerstone of the main uniqueness result of this paper (i.e., Theorem 1 in the sequel) and leads us naturally to recall the relationship between the rank of $A \odot B$ and identifiability of the CP model [41], i.e., CP is identifiable only if $A \odot B$ has full column rank. We introduce next a condition ensuring full column rank of $A \odot B$.

**Lemma 1:** Let $A \in \mathbb{C}^{M \times K}$ and $B \in \mathbb{C}^{P \times K}$ be two matrices consisting of nonzero columns. If either $\text{rank}(A) + k_B \geq K + 1$ or $\text{rank}(B) + k_A \geq K + 1$ holds, then $A \odot B$ is full column rank.

**Proof:** See Appendix A.

In [29], a similar lemma was presented but demanding $k_A + k_B \geq K + 1$. Comparatively, it is a milder condition posed here in Lemma 1. This lemma shows that a full column rank Khatri-Rao matrix can be generated with two rank-deficient matrices, which is also the basis of “signal decorrelation” techniques such as spatial averaging [42] or the polarization smoothing algorithm (PSA) [37]. This issue will be further discussed in Section IV. Generally, the necessary and sufficient uniqueness condition of Jiang and Sidiropoulos [39] is not easy to verify in practice. Towards this end, we provide a sufficient condition for identifiability of the CP model with a full column rank loading matrix (see also [35]) expressed in terms of rank and $k$-rank of the loading matrices.

**Theorem 1:** The CP model (5) is identifiable if the loading matrix $C$ is full column rank and the other two loading matrices $A$ and $B$ satisfy the conditions 1) $k_A, k_B \geq 2$, and 2) $\text{rank}(A) + k_B \geq K + 2$ or $\text{rank}(B) + k_A \geq K + 2$

(7)

**Proof:** See Appendix B.

The same condition was also proven independently by Stegeman [36], who found it sufficient for full column rank of the prespecified matrix $U$ and hence arrived at the assertion of uniqueness (see [36] for the details). This paper, however, provides (Appendix B) an alternative proof, using Lemma 1. It is also worth mentioning that the condition (7) is equivalent to Kruskal’s condition if $\text{rank}(A) = k_A$.

As shown in [36], the sufficient uniqueness condition of [39] based on the rank test of $U$ is more relaxed than that of Theorem 1. However, the goal of this paper, besides providing an identifiability condition efficient in use, is to find out the underlying link between the numbers of vector sensors and of the identifiable sources. Computing the prespecified matrix $U$ from the elements of $A$ and $B$ is intrinsically an element-wise nonlinear transformation. This may cause difficulties in tracking the contribution of a specific parameter, e.g., DOA, to the rank of this prespecified matrix. Hence, though Theorem 1 is more restrictive, it may provide useful perspectives on the uniqueness problem associated with the DOA estimation. To get more insights into the understanding of the two sufficient conditions for the particular case of the DOA identification problem, numerical simulations are conducted in the next section.

IV. APPLICATIONS TO THE PROBLEM OF UNIQUENESS IN DOA ESTIMATION WITH A VECTOR SENSOR ARRAY

In the sequel, based on the uniqueness results presented in the previous section, we derive upper bounds on the number of identifiable source DOA’s using a vector sensor array. Let us return to the DOA estimation problem posed in Section II, and recall the CP data model (3). The question is, what is the maximum number of the sources, $K$, whose DOA’s can be uniquely identified from (3), without a priori knowledge on the exact DOA/polarization March 14, 2011 DRAFT
parameters of the signals. We restrict ourselves to the following two typical scenarios with respect to the level of correlation among the impinging signals. In the first scenario the signals are assumed to be uncorrelated or partially correlated, implying full column rank of the signal matrix \( S \). In the second scenario, fully correlated (coherent) signals are considered meaning that \( S \) is rank deficient. A natural assumption considered for both scenarios is \( \text{A1}) \ the \ DOA's \ of \ the \ signals \ are \ distinct. \) Unless otherwise mentioned, our study will focus on the complete \( (P = 6 \) components) vector sensors, but the obtained results apply with minor modifications to all types of trimmed vector sensors.

\[ \text{A. The case of non-coherent signals} \]

We examine firstly the important case of \( \text{A2}) \ uncorrelated \ or \ partially \ correlated \ signals. \) Thus, the matrix \( S \) in \( (3) \) is full column rank. It has been proven \[34\] that for a linear equally spaced (LES) array with \( M \geq K \) vector sensors the assumptions \( \text{A1}) \) and \( \text{A2}) \) are sufficient for the DOA identifiability of all the \( K \) signals.

However, the assumption \( M \geq K \) required in \[34\] is not always necessary. Herein we consider the DOA identifiability problem for a sensor array of \( M < K \), arbitrarily positioned, complete vector sensors. To tackle this, one needs to know the linear dependence pattern of the steering vectors \( a(\phi, \psi) \) of the corresponding scalar sensor array, and that of one vector sensor \( b(\phi, \psi, \alpha, \beta) \).

For \( a(\phi, \psi) \) we follow the common assumption that \( \text{A3}) \ the \ manifold \ of \ the \ corresponding \ scalar \ sensor \ array \ a(\phi, \psi) \ is \ free \ from \ rank-M \ ambiguity, \) meaning that \( k_A = M \). This assumption is justifiable in many cases, e.g., for an LES array \( a(\phi, \psi) \) is a Vandermonde vector. Under the given conditions, we can deduce from Theorem 1 an upper bound on the number \( K \) of identifiable sources using a vector sensor array

\[
K \leq M + \text{rank}(B) - 2.
\]  

(8)

For the same scenario, Kruskal’s condition \( (6) \) yields \( K \leq M + k_B - 2 \), an upper bound which is more restrictive than \( (8) \).

Regarding the manifold of one vector sensor \( b(\phi, \psi, \alpha, \beta) \), some interesting results can be found in \[24\] and \[25\]. For instance, the manifold of a complete vector sensor is free of rank-2 ambiguity \[24\], meaning \( k_B \geq 3 \). Moreover, higher ranks of \( B \) can be expected, depending on the polarization parameters of the signals, but in general it holds that \( \text{rank}(B) \geq 4 \) unless there exist four signals with identical ellipticity angles \[24\]. Particularly, when the ellipticity angles for four out of the \( K \) signals are identical and equal to \( \pi/4 \), i.e., the four are circularly polarized with the same spin, \( k_B = 3 \) but \( \text{rank}(B) \) is still likely to be 6.

The typical example mentioned above highlights the following two facts. The first one is that the number of signals that can be uniquely localized by a vector sensor array is underestimated if Kruskal’s condition is used and there is a potential increase by applying \( (8) \). The second is that the increase, however, can not be very significant for the vector sensors since the maximum of \( \text{rank}(B) \) is the number of components of a vector sensor, which is no greater than 6.
In order to get further insights on the link between Theorem 1 and DOA identifiability let us recall a result from [24], stating that for a vector sensor array, the ambiguities become inevitable when \( K \geq 3M \) signals are present. To sufficiently ensure uniqueness, our condition simply excludes all the possible values of \( K \), for which ambiguities may occur, despite how unlikely it does happen. Therefore, a large gap can be observed between the derived bound (8), sufficient for uniqueness in localizing \( K \) signals without knowing exactly the true parameters of the signals, and the bound necessary for uniqueness, \( K < 3M \). In between the two bounds there is a domain where a general uniqueness condition for our problem is still missing. From a practical point of view, a part of this domain can be explored on a case-by-case basis using the sufficient uniqueness condition of [39] based on the matrix \( U \). Fig. 2 illustrates a comparison between the two sufficient conditions for uniqueness. For the simulations, we used an \((M \times K)\) matrix \( A \) with rank\((A) = k_A = M \) and a \((6 \times K)\) matrix \( B \). We set \( k_B = 3 \) unless \( B \) becomes full column rank where \( k_B = \text{rank}(B) = K \). Meanwhile, observe that \( B \) can be a full column rank matrix only if \( K \leq 6 \). We plotted the condition of Theorem 1 and the condition based on the rank of \( U \) with respect to \( k_A (= \text{rank}(A)) \) and \( \text{rank}(B) \) for four different values of \( K \). The black region in Fig.2 corresponds to both uniqueness conditions being satisfied, the dark grey indicates that only the second condition (based on the matrix \( U \)) is satisfied while the clear shade of gray corresponds to values of \((k_A, \text{rank}(B))\) for which uniqueness cannot be assessed by either of these two conditions. The white area indicates cases that are impossible for a given \( K \). One can see that the uniqueness test based on the matrix \( U \) becomes interesting in practical situations that are not covered by Theorem 1, especially when the number of components (i.e., \( K \)) is high. Although the general value of this result cannot be assessed, simulations conducted with various realizations of \( A \) and \( B \) (e.g., the entries of \( A \) and \( B \) randomly drawn from a continuous distribution), yielded the same result each time.

B. The case of coherent signals

Usually, coherence among the signals occurs as a result of multipath propagation environment and results in proportional columns in the matrix \( S \) (Eq. (3)). This implies a maximum rank of \( S \) equal to \( K - 1 \) and \( k_S = 1 \). In spite of this, we can assume, without loss of generality, that both assumptions A1) and A3) still hold. In the presence of correlated sources, the eigenstructure-based DOA estimation techniques, such as MUSIC, encounter difficulties. To tackle this problem, “signal decorrelation” techniques, such as the polarization smoothing algorithm (PSA) [37], can be used for vector sensor arrays. The PSA can be regarded as forming a new signal matrix \( \tilde{S} \triangleq B \odot S \) such that \( \tilde{S} \) is full column rank. Following Lemma 1, \( \tilde{S} \) is ensured to be full column rank if

\[
K \leq k_B + \text{rank}(S) - 1.
\]

The newly obtained data structure can thus be expressed as \( \tilde{Y} \triangleq A\tilde{S}^T \). To implement MUSIC-like algorithms on the new data \( \tilde{Y} \), a fundamental requirement is \( M \geq K + 1 \), i.e., more vector sensors than sources are needed [37]. Combining this condition with (9), one obtains

\[
K \leq \min\{M, k_B + \text{rank}(S)\} - 1
\]
Using the CP formalism of the data processed by the PSA, we show next that (10) is not only an algorithm implementation requirement but rather a structural condition to guarantee source DOA’s identifiability. We only consider the LES array. As there are more vector sensors than sources one can divide the sensor array into $L = M - K + 1$ overlapping subarrays, each consisting of $K$ vector sensors. The vector sensors numbered $m = 1, \ldots, K$ form the first subarray, those numbered $m = 2, \ldots, K + 1$ form the second subarray, etc. Let $J_l$, $l = 1, \ldots, L$, be the $K \times M$ selection matrices such that $J_l \tilde{Y}$ generates a $K \times K$ matrix composed of the $K$ rows of $\tilde{Y}$ corresponding to the $l$th subarray. Similarly to [28], using the new formulation, the data can be rearranged as:

$$\tilde{Y} = \begin{bmatrix} J_1 \tilde{Y} \\ \vdots \\ J_{M-K+1} \tilde{Y} \end{bmatrix} = \begin{bmatrix} J_1 A \\ \vdots \\ J_{M-K+1} A \end{bmatrix} \tilde{S}^T \quad (11)$$

We shall also introduce the following two matrices: $H \triangleq J_1 A$ and $\tilde{A}$, an $L \times K$ matrix consisting of the first $L$ rows of $A$ only. Under the considered assumptions $A$, $H$ and $\tilde{A}$ are all Vandermonde matrices. The matrix pencil $J_l A = HD_l(\tilde{A})$ is thus obtained, allowing to express (11) as the CP model

$$\tilde{Y} = (\tilde{A} \odot H) \tilde{S}^T. \quad (12)$$

Let us investigate next identifiability of the system (12). $\tilde{S}$ is full column rank if (9) is satisfied. Furthermore, given their Vandermonde structure, we have $k_H = K$ and $k_{\tilde{A}} = L$. Since $M \geq K + 1$, then $L \geq 2$ and the condition of
Theorem 1, \( \text{rank}(\tilde{A}) + k_{\mathbf{H}} \geq K + 2 \), is satisfied. Thus, under the given assumptions, the condition (10) guarantees that the \( K \) coherent signals can be uniquely localized. It should be noticed that, since \( k_{\tilde{A}} = \text{rank}(\tilde{A}) \), the condition of Theorem 1 coincides with Kruskal’s condition; hence, the same conclusion can also be drawn using the latter. Particularly, in the special case where \( \text{rank}(\mathbf{S}) = 1 \), i.e., signals are all coherent, (10) reduces to

\[
K \leq \min(k_{\mathbf{B}}, M - 1),
\]

which recalls the result of [37] in the case of an LES array (with complete vector sensors): \( K \leq \min(6, M - 1) \). Nevertheless, the result of [37] is not sufficient to ensure uniqueness and it can be regarded as a derivation of (13) under the optimistic assumption that \( k_{\mathbf{B}} = 6 \). A counterexample to this assumption is that for a complete vector sensor (6 components), one can have \( k_{\mathbf{B}} = 3 \) if four of the sources are circularly polarized with the same spin direction [24]. In this case, using the PSA only may not be sufficient to ensure system identifiability.

V. CONCLUSIONS

In this paper we establish a link between uniqueness of the three-way CP model and identifiability of the polarized source mixture recorded on a vector sensor array. A sufficient condition, more relaxed than Kruskal’s, for uniqueness of the CP decomposition of a three-way array, when one of the loading matrices has full column rank, is also provided. The proof for this condition is based on a newly derived sufficient condition on full column rank of the Khatri-Rao product of two matrices. These new uniqueness results are applied to study identifiability of the DOA’s using a vector sensor array in the cases of uncorrelated/partially correlated and coherent signals.

Generally, the uniqueness conditions derived in the paper are still too strong but they provide some interesting insights on the analyzed problem. In future work we aim at relaxing these conditions by finding and excluding the real causes of the ambiguities through a more refined analysis of the underlying physical phenomena. Furthermore, the problem can also be addressed from a probabilistic perspective, following Wax and Ziskind [17] who derived a much weaker condition ensuring unique localization of multiple sources “almost surely” by a scalar sensor array. This provides another clue in pursuing a tighter bound on the number of sources that can be uniquely localized using a vector sensor array, for which De Lathauwer’s generic uniqueness condition [40], properly adapted, could be very useful.

APPENDIX A

PROOF OF LEMMA 1

We need the following lemma for the proof.

Lemma 2 (Sidiropoulos and Liu [43]): Let \( \mathbf{A} \) be an \( I \times K \) matrix, and \( \tilde{\mathbf{A}} \) be an \( I \times n \) matrix consisting of any \( n \) columns on \( \mathbf{A} \). Then \( \min(n, k_{\tilde{A}}) \leq k_{\tilde{A}} \leq n \).

Proof of Lemma 1: The proof is somewhat similar to the one in [26]. We prove the condition \( \text{rank}(\mathbf{A}) + k_{\mathbf{B}} \geq K + 1 \) by contradiction. Let us assume that \( \mathbf{A} \odot \mathbf{B} \) is rank deficient. Then, there exists a set of scalars \( c_1, c_2, \ldots, c_K \), not all of which are zero, such that the linear combination of the columns of \( \mathbf{A} \odot \mathbf{B} \) weighted by them is equal to
zero. Let \( n \) be the number of nonzero elements among \( \{c_k\}_{k=1}^{K} \); since none of the columns of \( \mathbf{A} \) or \( \mathbf{B} \) is zero, \( n \neq 1 \), therefore, we set \( n \geq 2 \). Suppose, without loss of generality, that \( c_1, \ldots, c_n \) are the \( n \) nonzero coefficients and \( \mathbf{c} = [c_1, c_2, \ldots, c_n]^T \). Moreover, denote by \( \tilde{\mathbf{A}} \in \mathbb{C}^{J \times n} \) and \( \tilde{\mathbf{B}} \in \mathbb{C}^{J \times n} \) the two matrices, consisting of the first \( n \) columns of \( \mathbf{A} \) and \( \mathbf{B} \), respectively, corresponding to the \( n \) nonzero coefficients. Then the assumption can be reformulated as \((\tilde{\mathbf{A}} \circ \tilde{\mathbf{B}}) \mathbf{c} = \mathbf{0}\) or, equivalently \( \tilde{\mathbf{B}}(\tilde{\mathbf{A}} \text{diag}(\mathbf{c}))^T = \mathbf{0}\), where \( \text{diag}(\mathbf{c}) \) is a diagonal matrix containing the elements of vector \( \mathbf{c} \) on its main diagonal. In that case the columns of the matrix \([\tilde{\mathbf{A}} \text{diag}(\mathbf{c})]^T\) all lie in the null space of \( \tilde{\mathbf{B}} \), denoted by \( \text{null}(\tilde{\mathbf{B}}) \). Let \( \text{span}(\cdot) \) denote the subspace spanned by the columns of a matrix. Then \( \text{span}([\tilde{\mathbf{A}} \text{diag}(\mathbf{c})]^T) \subseteq \text{null}(\tilde{\mathbf{B}}) \). As the coefficients \( c_1, \ldots, c_n \) and columns of \( \mathbf{A} \) are all nonzero, there is at least one nonzero vector in the null space, meaning that \( n > \text{rank}(\tilde{\mathbf{B}}). \) Thus, \( k_B < n \) and, using Lemma 2, one obtains \( k_B \leq k_{\tilde{B}} \leq \text{rank}(\tilde{\mathbf{B}}) \), which, together with the rank-nullity theorem: \( \text{rank}(\tilde{\mathbf{B}}) + \dim[\text{null}(\tilde{\mathbf{B}})] = n \), yields
\[
\dim[\text{null}(\tilde{\mathbf{B}})] \leq n - k_{\tilde{B}}.
\tag{A.1}
\]
Since \( \text{span}([\tilde{\mathbf{A}} \text{diag}(\mathbf{c})]^T) \subseteq \text{null}(\tilde{\mathbf{B}}) \) and \( \text{diag}(\mathbf{c}) \) is full rank, one obtains: \( \text{rank}([\tilde{\mathbf{A}} \text{diag}(\mathbf{c})]^T) = \text{rank}(\tilde{\mathbf{A}}) \leq \dim[\text{null}(\tilde{\mathbf{B}})] \). Using (A.1), the previous relation becomes:
\[
\text{rank}(\tilde{\mathbf{A}}) \leq \dim[\text{null}(\tilde{\mathbf{B}})] \leq n - k_{\tilde{B}}.
\tag{A.2}
\]
Since \( \mathbf{A} \) contains all the columns of \( \tilde{\mathbf{A}} \) as well as the other \( (K - n) \) columns, its rank can not exceed \( \text{rank}(\tilde{\mathbf{A}}) \) by more than \( (K - n) \), that is, \( \text{rank}(\mathbf{A}) \leq \text{rank}(\tilde{\mathbf{A}}) + K - n \leq (n - k_{\tilde{B}}) + K - n = K - k_B. \) Thus we obtain \( \text{rank}(\mathbf{A}) + k_B \leq K \), which contradicts the condition: \( \text{rank}(\mathbf{A}) + k_B \geq K + 1 \).

The alternative condition \( \text{rank}(\mathbf{B}) + k_A \geq K + 1 \) can be proved in a similar way by simply exchanging the roles of \( \mathbf{A} \) and \( \mathbf{B} \). Combining the two conditions, Lemma 1 follows. \( \square \)

**APPENDIX B**

**PROOF OF THEOREM 1**

The proof is based on Jiang and Sidiropoulos's condition [39], presented briefly in Section III. Once again the proof is conducted by contradiction. Assume the decomposition is not unique, meaning that there exist two vectors \( \tilde{\mathbf{a}} \) and \( \tilde{\mathbf{b}} \) such that their tensor product \( \tilde{\mathbf{a}} \circ \tilde{\mathbf{b}} \) equals the linear combination of the columns of \( \mathbf{A} \circ \mathbf{B} \) weighted by some coefficients \( c_1, \ldots, c_K \in \mathbb{C} \). We assume that the first \( n \geq 2 \) elements among \( c_1, c_2, \ldots, c_K \) are nonzero, and correspondingly, \( \tilde{\mathbf{A}} \in \mathbb{C}^{M \times n} \) and \( \tilde{\mathbf{B}} \in \mathbb{C}^{P \times n} \) are the two matrices consisting of the first \( n \) columns of \( \mathbf{A} \) and \( \mathbf{B} \). If we denote \( \mathbf{c} = [c_1, c_2, \ldots, c_n]^T \) the assumption can be expressed as
\[
(\tilde{\mathbf{A}} \circ \tilde{\mathbf{B}}) \mathbf{c} = \tilde{\mathbf{a}} \circ \tilde{\mathbf{b}} \triangleq \tilde{\mathbf{d}}.
\tag{B.1}
\]

Next we analyze (B.1) for the two possible cases \( \tilde{\mathbf{d}} = \mathbf{0} \) and \( \tilde{\mathbf{d}} \neq \mathbf{0} \). If \( \tilde{\mathbf{d}} = \mathbf{0} \) it means that \( \mathbf{A} \circ \mathbf{B} \) is not full column rank. In view of Lemma 1, this implies: \( \text{rank}(\mathbf{A}) + k_B < K + 1 \) which contradicts the hypothesis of Theorem 1. Let us investigate now the case \( \tilde{\mathbf{d}} \neq \mathbf{0} \). First, one should note that condition 2) implies \( k_B \geq 2 \). Equation (B.1) is equivalent to \( \tilde{\mathbf{B}} \tilde{\mathbf{b}}(\tilde{\mathbf{A}} \text{diag}(\mathbf{c}) - \tilde{\mathbf{a}})^T = \mathbf{0} \) implying
\[
\text{rank}((\tilde{\mathbf{A}} \text{diag}(\mathbf{c}) - \tilde{\mathbf{a}})^T) \leq \dim(\text{null}(\tilde{\mathbf{B}} \tilde{\mathbf{b}}))
\tag{B.2}
\]
and \( \mathbf{b} \in \text{span}(\mathbf{B}) \); hence \( \text{span}(\mathbf{B} \mathbf{b}) = \text{span}(\mathbf{B}) \) and \( \text{rank}(\mathbf{B} \mathbf{b}) = \text{rank}(\mathbf{B}) \). Let us consider the following two cases with respect to \( \mathbf{a} \). a) If \( \mathbf{a} \notin \text{span}(\mathbf{A}) \) then \( \text{rank}(\mathbf{A} \text{diag}(\mathbf{c}) - \mathbf{a}) = \text{rank}(\mathbf{A}) + 1 \). Using the rank-nullity theorem one gets \( \text{dim}(\text{null}(\mathbf{B} \mathbf{b})) = n + 1 - \text{rank}(\mathbf{B} \mathbf{b}) = n + 1 - \text{rank}(\mathbf{B}) \), which, together with (B.2) yields \( \text{rank}(\mathbf{A}) + 1 \leq n + 1 - \text{rank}(\mathbf{B}) \), or equivalently, \( \text{rank}(\mathbf{A}) \leq n - \text{rank}(\mathbf{B}) \leq n - k_B \) which is identical to (A.2). Using a similar argumentation as in Appendix A, it is straightforward to show that \( \text{rank}(\mathbf{A}) \leq K - k_B \), and hence \( \text{rank}(\mathbf{A}) + k_B \leq K \) which yields a contradiction. b) If \( \mathbf{a} \in \text{span}(\mathbf{A}) \), then \( \text{rank}(\mathbf{A} \text{diag}(\mathbf{c}) - \mathbf{a}) = \text{rank}(\mathbf{A}) \). Then we can deduce from (B.2) that

\[
\text{rank}(\mathbf{A}) \leq n + 1 - \text{rank}(\mathbf{B}). \tag{B.3}
\]

If \( \text{rank}(\mathbf{B}) = n \) then \( \text{rank}(\mathbf{A}) \leq 1 \) which contradicts the hypothesis \( k_A \geq 2 \). Consequently, \( \text{rank}(\mathbf{B}) < n \) which implies \( k_B < n \). Then, following Lemma 2 (see Appendix A), one obtains that \( k_B \leq \text{rank}(\mathbf{B}) \). Similar to Appendix A, by taking into consideration the rest \( K - n \) columns of \( \mathbf{A} \), it follows from (B.3) that \( \text{rank}(\mathbf{A}) + k_B \leq K + 1 \), contradicting the hypothesis and completing the proof of Theorem 1. □

ACKNOWLEDGMENT

The authors wish to thank Dr. A. Stegeman (University of Groningen, The Netherlands) for the insightful discussions on uniqueness of the CP model, and the anonymous reviewers for their constructive comments and suggestions. The first author is also grateful to the associate editor for his encouraging comments.

REFERENCES


