

- [8] M. Unser, "Sampling 50 years after Shannon," *Proc. IEEE*, vol. 88, no. 4, pp. 569–587, Apr. 2000.
- [9] M. Unser, "Splines: A perfect fit for signal and image processing," *IEEE Signal Process. Mag.*, vol. 16, no. 6, pp. 22–38, Nov. 1999.
- [10] M. Unser and T. Blu, "Generalized smoothing splines and the optimal discretization of the wiener filter," *IEEE Trans. Signal Process.*, vol. 53, no. 6, pp. 2146–2159, Jun. 2005.
- [11] M. Unser, "Cardinal exponential splines: Part II—Think analog, act digital," *IEEE Trans. Signal Process.*, vol. 53, no. 4, pp. 1439–1449, Apr. 2005.
- [12] M. Unser and T. Blu, "Self-similarity: Part I—Splines and operators," *IEEE Trans. Signal Process.*, vol. 55, no. 4, pp. 1352–1363, Apr. 2007.
- [13] A. Björck and G. H. Golub, "Numerical methods for computing the angles between linear subspaces," *Math. Comput.*, pp. 579–594, 1973.

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 (DOAs) 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 DOAs 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—CANDECOMP/PARAFAC uniqueness, identifiability, polarization, vector sensor array processing.

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 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 the ones 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 DOAs 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 DOAs of the superimposed signals.

In this context, a problem of primary importance is to find the conditions under which the DOAs 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 [37], 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.

Manuscript received July 13, 2010; revised November 18, 2010 and February 12, 2011; accepted March 01, 2011. Date of publication March 22, 2011; date of current version June 15, 2011. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Xavier Mestre. 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 High Technology Research and Development Program ("863" Program) of China under Grant 2009AA011502. Part of the work (Theorem 1 and the proof) was presented at the Twenty-Second GRETSI Symposium on Signal and Image Processing, Dijon, France, September 8–11, 2009. The work of X. Guo for this paper was performed while the author was at CRAN.

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, F-54506 Vandœuvre-lès-Nancy, France (e-mail: xijing.guo@cran.uhp-nancy.fr).

S. Miron and D. Brie are with Centre de Recherche en Automatique de Nancy (CRAN), Nancy-Université, CNRS, F-54506 Vandœuvre-lès-Nancy, France (e-mail: 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.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TSP.2011.2131134

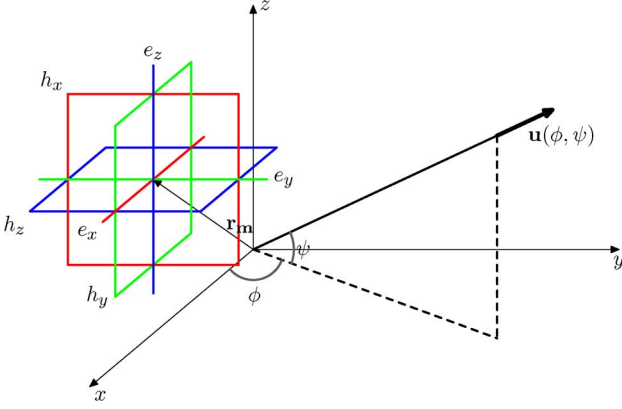


Fig. 1. Geometry of a vector sensor: $\{e_x, e_y, e_z\}$ denotes the dipole triad and $\{h_x, h_y, h_z\}$ the loop triad.

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]

$$\mathbf{b}(\phi, \psi, \alpha, \beta) \triangleq \underbrace{\begin{bmatrix} -\sin \phi & -\cos \phi \sin \psi \\ \cos \phi & -\sin \phi \sin \psi \\ 0 & \cos \psi \\ -\cos \phi \sin \psi & \sin \phi \\ -\sin \phi \sin \psi & -\cos \phi \\ \cos \psi & 0 \end{bmatrix}}_{\mathbf{\Theta}(\phi, \psi)} \times \underbrace{\begin{bmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{bmatrix} \begin{bmatrix} \cos \beta \\ j \sin \beta \end{bmatrix}}_{\mathbf{P}(\alpha, \beta)} \quad (1)$$

for an incident signal with DOA (ϕ, ψ) and polarization (α, β) . In (1), $\phi \in [0, 2\pi]$ denotes the azimuth angle, $\psi \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ is the elevation angle, $\alpha \in (-\frac{\pi}{2}, \frac{\pi}{2}]$ is the orientation angle of the polarization ellipse and $\beta \in [-\frac{\pi}{4}, \frac{\pi}{4}]$ is the ellipticity angle. Throughout the paper, we assume that the DOAs and the polarizations are *unknown but deterministic*.

We denote by a set of three-dimensional vectors $\{\mathbf{r}_m\}_{m=1}^M$, the M observation points where the vector sensors are positioned in the reference frame, and by $\mathbf{u}(\phi, \psi) = [\cos \phi \cos \psi \ \sin \phi \cos \psi \ \sin \psi]^T$ the unit Poynting vector in the source direction. The expression for the phase shifts induced by the displacements of the m th ($m = 1, \dots, M$) vector sensor relative to the reference one is then given by $a_m(\phi, \psi) =$

$\exp\left\{j2\pi \mathbf{r}_m^T \frac{\mathbf{u}(\phi, \psi)}{\lambda}\right\}$, where λ is the wavelength. Furthermore, define $\mathbf{a}(\phi, \psi) = [a_1(\phi, \psi), \dots, a_M(\phi, \psi)]^T$, 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] $\mathbf{d}(\phi, \psi, \alpha, \beta) = \mathbf{a}(\phi, \psi) \otimes \mathbf{b}(\phi, \psi, \alpha, \beta)$ where \otimes denotes the Kronecker product. Now suppose that the k th signal impinges on the array from a direction (ϕ_k, ψ_k) , with the polarization parameterized by (α_k, β_k) . Inductively, the output of the array at time t , i.e., $\mathbf{y}(t)$, is obtained by summing up the contributions from all the K incident signals $s_k(t)$:

$$\mathbf{y}(t) = \sum_{k=1}^K s_k(t) \mathbf{d}(\phi_k, \psi_k, \alpha_k, \beta_k) = \mathbf{D} \mathbf{s}(t) \quad (2)$$

where $\mathbf{D} = [\mathbf{d}(\phi_1, \psi_1, \alpha_1, \beta_1), \dots, \mathbf{d}(\phi_K, \psi_K, \alpha_K, \beta_K)]$ is a $6M \times K$ matrix, and $\mathbf{s}(t) = [s_1(t), \dots, s_K(t)]^T$ 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 $\mathbf{A} \triangleq [\mathbf{a}(\phi_1, \psi_1), \dots, \mathbf{a}(\phi_K, \psi_K)]$, and $\mathbf{B} \triangleq [\mathbf{b}(\phi_1, \psi_1, \alpha_1, \beta_1), \dots, \mathbf{b}(\phi_K, \psi_K, \alpha_K, \beta_K)]$. Then, one observes $\mathbf{D} = \mathbf{A} \odot \mathbf{B}$ 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 $\mathbf{Y} = [\mathbf{y}(t_1), \dots, \mathbf{y}(t_N)] = \mathbf{D} \mathbf{S}^T$, where $\mathbf{S} \triangleq [\mathbf{s}(t_1), \dots, \mathbf{s}(t_N)]^T$. Observing the multiple invariances of the data model, \mathbf{Y} can also be expressed as

$$\mathbf{Y} \triangleq \begin{bmatrix} \mathbf{Y}_1 \\ \vdots \\ \mathbf{Y}_M \end{bmatrix} = \begin{bmatrix} \mathbf{B} \mathbf{D}_1(\mathbf{A}) \\ \vdots \\ \mathbf{B} \mathbf{D}_M(\mathbf{A}) \end{bmatrix} \mathbf{S}^T = (\mathbf{A} \odot \mathbf{B}) \mathbf{S}^T \quad (3)$$

where $\mathbf{D}_m(\mathbf{A}) = \text{diag}(a_{m1}, \dots, a_{mK})$ denotes the diagonal matrix with the m th row of $\mathbf{A} = [a_{mk}]_{M \times K}$ as its diagonal, and $\mathbf{Y}_m = \mathbf{B} \mathbf{D}_m(\mathbf{A}) \mathbf{S}^T$. 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*) \mathcal{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 \mathbf{a} , \mathbf{b} and \mathbf{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 \mathcal{X} is rank- K if K is the minimum number of rank-1 tensors in the CP decomposition

$$x_{mpn} = \sum_{k=1}^K a_{mk} b_{pk} c_{nk} \quad (4)$$

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

$$\mathbf{X} \triangleq \begin{bmatrix} \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_M \end{bmatrix} = (\mathbf{A} \odot \mathbf{B}) \mathbf{C}^T. \quad (5)$$

The CP decomposition of a tensor presents interesting uniqueness¹ 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 $\mathbf{A} \in \mathbb{C}^{M \times K}$ equals k_A , if k_A is the maximum number of ℓ such that every ℓ columns of \mathbf{A} are linearly independent.

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

Kruskal’s condition [33]

$$k_A + k_B + k_C \geq 2K + 2 \quad (6)$$

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 \mathbf{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 follows:

The CP decomposition is unique if *none* of the nontrivial² linear combinations of the columns of $\mathbf{A} \odot \mathbf{B}$ 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 \mathbf{C} was also provided in [39] (see also [36]). It depends on a prespecified $M^2 P^2 \times \frac{K(K-1)}{2}$ matrix \mathbf{U} whose elements are determined by the product of the second-order minors of \mathbf{A} and \mathbf{B} (see [39] for the details on how to generate \mathbf{U}). Jiang and Sidiropoulos [39] have proven that if \mathbf{U} has full column rank then identifiability of the model is guaranteed. The same condition was derived independently by De Lathauwer in [40].

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 $\mathbf{A} \odot \mathbf{B}$ and identifiability of the CP model [41], i.e., CP is identifiable only if $\mathbf{A} \odot \mathbf{B}$ has full column rank. We introduce next a condition ensuring full column rank of $\mathbf{A} \odot \mathbf{B}$.

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

Proof: See Appendix A. \square

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.

¹Here, by uniqueness, we mean “essential uniqueness,” that is, if another set of matrices ($\tilde{\mathbf{A}}$, $\tilde{\mathbf{B}}$, and $\tilde{\mathbf{C}}$) also satisfy (5) exactly, there exists a permutation matrix Π and three invertible diagonal scaling matrices Δ_1 , Δ_2 , and Δ_3 satisfying $\Delta_1 \Delta_2 \Delta_3 = \mathbf{I}$, where \mathbf{I} is the identity matrix, such that $\tilde{\mathbf{A}} = \mathbf{A} \Pi \Delta_1$, $\tilde{\mathbf{B}} = \mathbf{B} \Pi \Delta_2$, $\tilde{\mathbf{C}} = \mathbf{C} \Pi \Delta_3$.

²The *nontrivial* linear combination is referred to as a combination involving at least two vectors.

³The tensor product of two vectors \mathbf{a} and \mathbf{b} is given by their element-wise product (Kronecker product) $\mathbf{a} \otimes \mathbf{b}$.

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

$$\text{rank}(\mathbf{A}) + k_B \geq K + 2 \quad \text{or} \quad \text{rank}(\mathbf{B}) + k_A \geq K + 2. \quad (7)$$

Proof: See Appendix B. \square

The same condition was also proven independently by Stegeman [36], who found it *sufficient* for full column rank of the prespecified matrix \mathbf{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 $\text{rank}(\mathbf{A}) = k_A$.

As shown in [36], the sufficient uniqueness condition of [39] based on the rank test of \mathbf{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 \mathbf{U} from the elements of \mathbf{A} and \mathbf{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 DOAs 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 DOAs can be uniquely identified from (3), without *a priori* knowledge on the exact DOA/polarization 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 \mathbf{S} . In the second scenario, fully correlated (coherent) signals are considered meaning that \mathbf{S} is rank deficient. A natural assumption considered for both scenarios is A1) *the DOAs 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.

A. The Case of Noncoherent Signals

We examine firstly the important case of A2) *uncorrelated or partially correlated signals*. Thus, the matrix \mathbf{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 A1) and 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 $\mathbf{a}(\phi, \psi)$ of the corresponding scalar sensor array, and that of one vector sensor $\mathbf{b}(\phi, \psi, \alpha, \beta)$.

For $\mathbf{a}(\phi, \psi)$ we follow the common assumption that A3) *the manifold of the corresponding scalar sensor array $\mathbf{a}(\phi, \psi)$ is free from rank- M ambiguity*, meaning that $k_A = M$. This assumption is justifiable in many cases, e.g., for an LES array $\mathbf{a}(\phi, \psi)$ is a Vandermonde vector.

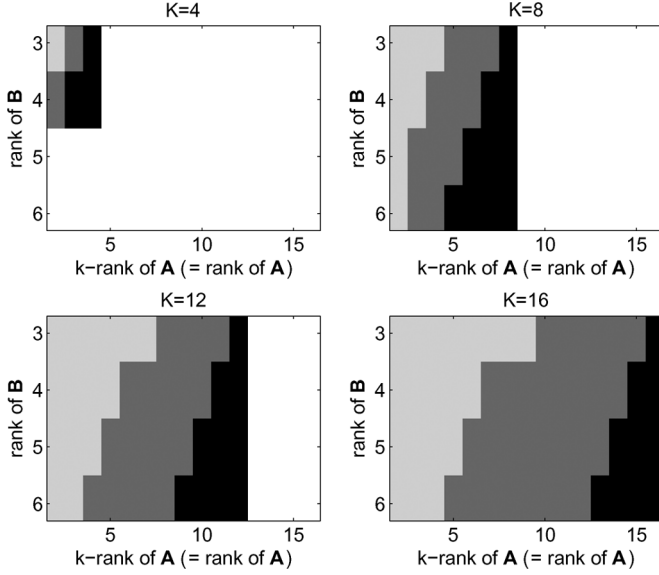


Fig. 2. A comparison between the sufficient conditions for DOA identifiability given by Theorem 1 and the one based on matrix \mathbf{U} [39]. The black region corresponds to both conditions being satisfied, the dark gray indicates that only the condition of [39] is verified and the clear shade of gray indicates that neither of the two uniqueness conditions works.

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

$$K \leq M + \text{rank}(\mathbf{B}) - 2. \quad (8)$$

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 $\mathbf{b}(\phi, \psi, \alpha, \beta)$, 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 \mathbf{B} can be expected, depending on the polarization parameters of the signals, but in general it holds that $\text{rank}(\mathbf{B}) \geq 4$ 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 $\frac{\pi}{4}$, i.e., the four are circularly polarized with the same spin, $k_B = 3$ but $\text{rank}(\mathbf{B})$ 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, cannot be very significant for the vector sensors since the maximum of $\text{rank}(\mathbf{B})$ 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 \mathbf{U} . Fig. 2 illustrates a comparison between the two sufficient conditions for uniqueness. For the simulations, we used an $(M \times K)$ matrix \mathbf{A} with

$\text{rank}(\mathbf{A}) = k_A = M$ and a $(6 \times K)$ matrix \mathbf{B} . We set $k_B = 3$ unless \mathbf{B} becomes full column rank where $k_B = \text{rank}(\mathbf{B}) = K$. Meanwhile, observe that \mathbf{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 \mathbf{U} with respect to $k_A (= \text{rank}(\mathbf{A}))$ and $\text{rank}(\mathbf{B})$ for four different values of K . The black region in Fig. 2 corresponds to both uniqueness conditions being satisfied, the dark gray indicates that only the second condition (based on the matrix \mathbf{U}) is satisfied while the clear shade of gray corresponds to values of $(k_A, \text{rank}(\mathbf{B}))$ 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 \mathbf{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 \mathbf{A} and \mathbf{B} (e.g., the entries of \mathbf{A} and \mathbf{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 \mathbf{S} (3). This implies a maximum rank of \mathbf{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 $\check{\mathbf{S}} \triangleq \mathbf{B} \odot \mathbf{S}$ such that $\check{\mathbf{S}}$ is full column rank. Following Lemma 1, $\check{\mathbf{S}}$ is ensured to be full column rank if

$$K \leq k_B + \text{rank}(\mathbf{S}) - 1. \quad (9)$$

The newly obtained data structure can thus be expressed as $\check{\mathbf{Y}} \triangleq \mathbf{A} \check{\mathbf{S}}^T$. To implement MUSIC-like algorithms on the new data $\check{\mathbf{Y}}$, 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

$$K \leq \min \{M, k_B + \text{rank}(\mathbf{S})\} - 1. \quad (10)$$

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 DOAs 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 $m = 1, \dots, K$ form the first subarray, those numbered $m = 2, \dots, K + 1$ form the second subarray, etc. Let \mathbf{J}_l , $l = 1, \dots, L$, be the $K \times M$ selection matrices such that $\mathbf{J}_l \check{\mathbf{Y}}$ generates a $K \times K$ matrix composed of the K rows of $\check{\mathbf{Y}}$ corresponding to the l th subarray. Similarly to [28], using the new formulation, the data can be rearranged as

$$\check{\mathbf{Y}} = \begin{bmatrix} \mathbf{J}_1 \check{\mathbf{Y}} \\ \vdots \\ \mathbf{J}_{M-K+1} \check{\mathbf{Y}} \end{bmatrix} = \begin{bmatrix} \mathbf{J}_1 \mathbf{A} \\ \vdots \\ \mathbf{J}_{M-K+1} \mathbf{A} \end{bmatrix} \check{\mathbf{S}}^T. \quad (11)$$

We shall also introduce the following two matrices: $\check{\mathbf{H}} \triangleq \mathbf{J}_1 \mathbf{A}$ and $\check{\mathbf{A}}$, an $L \times K$ matrix consisting of the first L rows of \mathbf{A} only. Under the considered assumptions \mathbf{A} , $\check{\mathbf{H}}$ and $\check{\mathbf{A}}$ are all Vandermonde matrices. The matrix pencil $\mathbf{J}_1 \mathbf{A} = \check{\mathbf{H}} \mathbf{D}_1(\check{\mathbf{A}})$ is thus obtained, allowing to express (11) as the CP model

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

Let us investigate next identifiability of the system (12). $\tilde{\mathbf{S}}$ is full column rank if (9) is satisfied. Furthermore, given their Vandermonde structure, we have $k_{\mathbf{H}} = K$ and $k_{\tilde{\mathbf{A}}} = L$. Since $M \geq K + 1$, then $L \geq 2$ and the condition of Theorem 1, $\text{rank}(\tilde{\mathbf{A}}) + k_{\mathbf{H}} \geq K + 2$, 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 $k_{\tilde{\mathbf{A}}} = \text{rank}(\tilde{\mathbf{A}})$, 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

$$K \leq \min(k_{\mathbf{B}}, M - 1), \quad (13)$$

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 (six 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. CONCLUSION

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 DOAs 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

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 $\min(n, k_{\mathbf{A}}) \leq k_{\tilde{\mathbf{A}}} \leq n$.

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, \dots, 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, \dots, c_n are the n nonzero coefficients and $\mathbf{c} = [c_1, c_2, \dots, c_n]^T$. Moreover, denote by $\tilde{\mathbf{A}} \in \mathbb{C}^{I \times n}$ 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 $(\tilde{\mathbf{A}} \odot \tilde{\mathbf{B}})\mathbf{c} = \mathbf{0}$

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, \dots, 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, $k_{\tilde{\mathbf{B}}} < n$ and, using Lemma 2, one obtains $k_{\mathbf{B}} \leq k_{\tilde{\mathbf{B}}} \leq \text{rank}(\tilde{\mathbf{B}})$, which, together with the rank-nullity theorem: $\text{rank}(\tilde{\mathbf{B}}) + \dim[\text{null}(\tilde{\mathbf{B}})] = n$, yields

$$\dim[\text{null}(\tilde{\mathbf{B}})] \leq n - k_{\mathbf{B}}. \quad (A1)$$

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 (A1), the previous relation becomes

$$\text{rank}(\tilde{\mathbf{A}}) \leq \dim[\text{null}(\tilde{\mathbf{B}})] \leq n - k_{\mathbf{B}}. \quad (A2)$$

Since \mathbf{A} contains all the columns of $\tilde{\mathbf{A}}$ as well as the other $(K - n)$ columns, its rank cannot exceed $\text{rank}(\tilde{\mathbf{A}})$ by more than $(K - n)$, that is, $\text{rank}(\mathbf{A}) \leq \text{rank}(\tilde{\mathbf{A}}) + K - n \leq (n - k_{\mathbf{B}}) + K - n = K - k_{\mathbf{B}}$. Thus, we obtain $\text{rank}(\mathbf{A}) + k_{\mathbf{B}} \leq K$, which contradicts the condition: $\text{rank}(\mathbf{A}) + k_{\mathbf{B}} \geq K + 1$.

The alternative condition $\text{rank}(\mathbf{B}) + k_{\mathbf{A}} \geq K + 1$ 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 II. Once again the proof is conducted by contradiction. Assume the decomposition is not unique, meaning that there exist two vectors $\bar{\mathbf{a}}$ and $\bar{\mathbf{b}}$ such that their tensor product $\bar{\mathbf{a}} \otimes \bar{\mathbf{b}}$ equals the linear combination of the columns of $\mathbf{A} \odot \mathbf{B}$ weighted by some coefficients $c_1, \dots, c_K \in \mathbb{C}$. We assume that the first $n \geq 2$ elements among c_1, c_2, \dots, 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 $\mathbf{c} = [c_1, c_2, \dots, c_n]^T$ the assumption can be expressed as

$$(\tilde{\mathbf{A}} \odot \tilde{\mathbf{B}})\mathbf{c} = \bar{\mathbf{a}} \otimes \bar{\mathbf{b}} \triangleq \bar{\mathbf{d}}. \quad (B1)$$

Next we analyze (B1) for the two possible cases $\bar{\mathbf{d}} = \mathbf{0}$ and $\bar{\mathbf{d}} \neq \mathbf{0}$. If $\bar{\mathbf{d}} = \mathbf{0}$ it means that $\mathbf{A} \odot \mathbf{B}$ is not full column rank. In view of Lemma 1, this implies: $\text{rank}(\mathbf{A}) + k_{\mathbf{B}} < K + 1$ which contradicts the hypothesis of Theorem 1. Let us investigate now the case $\bar{\mathbf{d}} \neq \mathbf{0}$. First, one should note that condition 2) implies $k_{\mathbf{B}} \geq 2$. Equation (B1)

is equivalent to $[\tilde{\mathbf{B}}; \bar{\mathbf{b}}][\tilde{\mathbf{A}}\text{diag}(\mathbf{c}) - \bar{\mathbf{a}}]^T = \mathbf{0}$ implying

$$\text{rank} \left(\begin{bmatrix} \tilde{\mathbf{A}}\text{diag}(\mathbf{c}) - \bar{\mathbf{a}} \end{bmatrix}^T \right) \leq \dim \left(\text{null}[\tilde{\mathbf{B}}; \bar{\mathbf{b}}] \right) \quad (B2)$$

and $\bar{\mathbf{b}} \in \text{span}(\tilde{\mathbf{B}})$; hence $\text{span}([\tilde{\mathbf{B}}; \bar{\mathbf{b}}]) = \text{span}(\tilde{\mathbf{B}})$ and $\text{rank}([\tilde{\mathbf{B}}; \bar{\mathbf{b}}]) = \text{rank}(\tilde{\mathbf{B}})$. Let us consider the following two cases with respect to $\bar{\mathbf{a}}$. a) If $\bar{\mathbf{a}} \notin \text{span}(\tilde{\mathbf{A}})$ then $\text{rank}[\tilde{\mathbf{A}}\text{diag}(\mathbf{c}) - \bar{\mathbf{a}}] = \text{rank}(\tilde{\mathbf{A}}) + 1$. Using the rank-nullity theorem one gets $\dim(\text{null}[\tilde{\mathbf{B}}; \bar{\mathbf{b}}]) = n + 1 - \text{rank}([\tilde{\mathbf{B}}; \bar{\mathbf{b}}]) = n + 1 - \text{rank}(\tilde{\mathbf{B}})$, which, together with (B2) yields $\text{rank}(\tilde{\mathbf{A}}) + 1 \leq n + 1 - \text{rank}(\tilde{\mathbf{B}})$, or equivalently, $\text{rank}(\tilde{\mathbf{A}}) \leq n - \text{rank}(\tilde{\mathbf{B}}) \leq n - k_{\tilde{\mathbf{B}}}$ which is identical to (A2). Using a similar argumentation as in Appendix A, it is straightforward to show that $\text{rank}(\mathbf{A}) \leq K - k_{\mathbf{B}}$, and hence

$\text{rank}(\mathbf{A}) + k_B \leq K$ which yields a contradiction. b) If $\bar{\mathbf{a}} \in \text{span}(\tilde{\mathbf{A}})$, then $\text{rank}[\tilde{\mathbf{A}}\text{diag}(\mathbf{c})^T - \bar{\mathbf{a}}] = \text{rank}(\tilde{\mathbf{A}})$. Then we can deduce from (B2) that

$$\text{rank}(\tilde{\mathbf{A}}) \leq n + 1 - \text{rank}(\tilde{\mathbf{B}}). \quad (\text{B3})$$

If $\text{rank}(\tilde{\mathbf{B}}) = n$ then $\text{rank}(\tilde{\mathbf{A}}) \leq 1$ which contradicts the hypothesis $k_A \geq 2$. Consequently, $\text{rank}(\tilde{\mathbf{B}}) < n$ which implies $k_B < n$. Then, following Lemma 2 (see Appendix A), one obtains that $k_B \leq \text{rank}(\tilde{\mathbf{B}})$. Similar to Appendix A, by taking into consideration the rest $K - n$ columns of \mathbf{A} , it follows from (B3) that $\text{rank}(\mathbf{A}) + k_B \leq K + 1$, contradicting the hypothesis and completing the proof of Theorem 1. \square

ACKNOWLEDGMENT

The authors would like 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

- [1] A. Nehorai and E. Paldi, "Vector-sensor array processing for electromagnetic source localization," *IEEE Trans. Signal Process.*, vol. 42, no. 2, pp. 376–398, Feb. 1994.
- [2] A. J. Weiss and B. Friedlander, "Direction finding for diversely polarized signals using polynomial rooting," *IEEE Trans. Signal Process.*, vol. 41, no. 5, pp. 1893–1905, May 1993.
- [3] A. L. Swindlehurst and M. Viberg, "Subspace fitting with diversely polarized antenna arrays," *IEEE Trans. Antennas Propag.*, vol. 41, no. 12, pp. 1687–1694, Dec. 1993.
- [4] J. Li, "Direction and polarization estimation using arrays with small loops and short dipoles," *IEEE Trans. Antennas Propag.*, vol. 41, no. 3, pp. 379–387, Mar. 1993.
- [5] K. T. Wong and M. D. Zoltowski, "Uni-vector-sensor ESPRIT for multiresource azimuth, elevation, and polarization estimation," *IEEE Trans. Antennas Propag.*, vol. 45, no. 10, pp. 1467–1474, Oct. 1997.
- [6] K. T. Wong and M. D. Zoltowski, "Closed-form direction finding and polarization estimation with arbitrarily spaced electromagnetic vector-sensors at unknown locations," *IEEE Trans. Antennas Propag.*, vol. 48, no. 5, pp. 671–681, May 2000.
- [7] K.-C. Ho, K.-C. Tan, and A. Nehorai, "Estimating directions of arrival of completely and incompletely polarized signals with electromagnetic vector sensors," *IEEE Trans. Signal Process.*, vol. 47, no. 10, pp. 2845–2852, Oct. 1999.
- [8] M. D. Zoltowski and K. T. Wong, "ESPRIT-based 2-D direction finding with a sparse uniform array of electromagnetic vector sensors," *IEEE Trans. Signal Process.*, vol. 48, no. 8, pp. 2195–2204, Aug. 2000.
- [9] M. D. Zoltowski and K. T. Wong, "Closed-form eigenstructure-based direction finding using arbitrary but identical subarrays on a sparse uniform cartesian array grid," *IEEE Trans. Signal Process.*, vol. 48, no. 8, pp. 2205–2210, Aug. 2000.
- [10] R. O. Schmidt, "A signal subspace approach to multiple emitter location and spectral estimation," Ph.D. dissertation, Elect. Eng. Dept., Stanford Univ., Stanford, CA, 1981.
- [11] K. T. Wong and M. D. Zoltowski, "Self-initiating MUSIC-based direction finding and polarization estimation in spatio-polarizational beamspace," *IEEE Trans. Antennas Propag.*, vol. 48, no. 8, pp. 1235–1245, Aug. 2000.
- [12] K. T. Wong, L. Li, and M. D. Zoltowski, "Root-MUSIC-based direction-finding and polarization estimation using diversely polarized possibly collocated antennas," *IEEE Antennas Wireless Propag. Lett.*, vol. 3, pp. 129–132, Aug. 2004.
- [13] S. Miron, N. Le Bihan, and J. I. Mars, "Quaternion MUSIC for vector-sensor array processing," *IEEE Trans. Signal Process.*, vol. 54, no. 4, pp. 1218–1229, Apr. 2006.
- [14] P. Chevalier, A. Ferréol, L. Albera, and G. Birot, "Higher order direction finding from arrays with diversely polarized antennas: The PD-2q-MUSIC algorithms," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5337–5350, Nov. 2007.
- [15] Y. Zhang, B. A. Obeidat, and M. G. Amin, "Spatial polarimetric time-frequency distributions for direction-of-arrival estimations," *IEEE Trans. Signal Process.*, vol. 54, no. 4, pp. 1327–1340, Apr. 2006.
- [16] X. Gong, Z. Liu, Y. Xu, and M. I. Ahmad, "Direction-of-arrival estimation via twofold mode-projection," *Signal Process.*, vol. 89, no. 5, pp. 831–842, May 2009.
- [17] M. Wax and I. Ziskind, "On unique localization of multiple sources by passive sensor arrays," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 996–1000, Jul. 1989.
- [18] B. Hochwald and A. Nehorai, "Identifiability in array processing models with vector-sensor applications," *IEEE Trans. Signal Process.*, vol. 44, no. 1, pp. 83–95, Jan. 1996.
- [19] K.-C. Ho, K.-C. Tan, and W. Ser, "An investigation on number of signals whose directions-of-arrival are uniquely determinable with an electromagnetic vector sensor," *Signal Process.*, vol. 47, no. 1, pp. 41–54, Nov. 1995.
- [20] J. T.-H. Lo and S. L. Marple, Jr., "Observability conditions for multiple signal direction finding and array sensor localization," *IEEE Trans. Signal Process.*, vol. 40, no. 11, pp. 2641–2650, Nov. 1992.
- [21] K.-C. Tan, S. S. Goh, and E.-C. Tan, "A study of the rank-ambiguity issues in direction-of-arrival estimation," *IEEE Trans. Signal Process.*, vol. 44, no. 4, pp. 880–887, Apr. 1996.
- [22] A. Manikas, H. R. Karimi, and I. Dacos, "Study of the detection and resolution capabilities of a one-dimensional array of sensors by using differential geometry," *Proc. Inst. Electr. Eng.—Radar, Sonar Navig.*, vol. 141, no. 2, pp. 83–92, Apr. 1994.
- [23] A. Manikas, A. Sleiman, and I. Dacos, "Manifold studies of nonlinear antenna array geometries," *IEEE Trans. Signal Process.*, vol. 49, no. 3, pp. 497–506, Mar. 2001.
- [24] K.-C. Tan, K.-C. Ho, and A. Nehorai, "Linear independence of steering vectors of an electromagnetic vector sensor," *IEEE Trans. Signal Process.*, vol. 44, no. 12, pp. 3099–3107, Dec. 1996.
- [25] K.-C. Ho, K.-C. Tan, and B. T. G. Tan, "Linear dependence of steering vectors associated with tripole arrays," *IEEE Trans. Antennas Propag.*, vol. 46, no. 11, pp. 1705–1711, Nov. 1998.
- [26] K.-C. Tan, K.-C. Ho, and A. Nehorai, "Uniqueness study of measurements obtainable with arrays of electromagnetic vector sensors," *IEEE Trans. Signal Process.*, vol. 44, no. 4, pp. 1036–1039, Apr. 1996.
- [27] Y. Xu, Z. Liu, K. T. Wong, and J. Cao, "Virtual-manifold ambiguity in HOS-based direction-finding with electromagnetic vector-sensors," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 44, no. 4, pp. 1291–1308, Oct. 2008.
- [28] A. L. Swindlehurst, B. Ottersten, R. Roy, and T. Kailath, "Multiple invariance ESPRIT," *IEEE Trans. Signal Process.*, vol. 40, no. 4, pp. 867–881, Apr. 1992.
- [29] N. D. Sidiropoulos, R. Bro, and G. B. Giannakis, "Parallel factor analysis in sensor array processing," *IEEE Trans. Signal Process.*, vol. 48, no. 8, pp. 2377–2388, Aug. 2000.
- [30] J. D. Carroll and J.-J. Chang, "Analysis of individual differences in multidimensional scaling via an N-way generalization of 'Eckart-Young' decomposition," *Psychometrika*, vol. 35, no. 3, pp. 283–319, Sep. 1970.
- [31] R. A. Harshman, "Foundations of the PARAFAC procedure: Model and conditions for an 'explanatory' multi-mode factor analysis," *UCLA Working Papers Phonetics*, vol. 16, pp. 1–84, Dec. 1970.
- [32] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, Sep. 2009.
- [33] J. B. Kruskal, "Three-way arrays: Rank and uniqueness of trilinear decompositions with application to arithmetic complexity and statistics," *Linear Algebra Appl.*, vol. 18, no. 2, pp. 95–138, 1977.
- [34] X. Guo, S. Miron, and D. Brie, "Identifiability of the PARAFAC model for polarized source mixture on a vector sensor array," in *Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Las Vegas, NV, Mar. 31–Apr. 4, 2008, pp. 2401–2404.
- [35] X. Guo, S. Miron, and D. Brie, "Identifiabilité du mélange de sources polarisées: le cas où le nombre de capteurs est inférieur au nombre de sources," presented at the GRETSI, (in French) Dijon, France, Sep. 2009.
- [36] A. Stegeman, "On uniqueness conditions for CANDECOMP/PARAFAC and INDSCAL with full column rank in one mode," *Linear Algebra Appl.*, vol. 431, no. 1–2, pp. 211–227, Jul. 2009.
- [37] D. Rahamim, J. Tabrikian, and R. Shavit, "Source localization using vector sensor array in a multipath environment," *IEEE Trans. Signal Process.*, vol. 52, no. 11, pp. 3096–3103, Nov. 2004.
- [38] J. M. F. ten Berge and N. D. Sidiropoulos, "On uniqueness in CANDECOMP/PARAFAC," *Psychometrika*, vol. 67, no. 3, pp. 399–409, Sep. 2002.

- [39] T. Jiang and N. D. Sidiropoulos, "Kruskal's permutation lemma and the identification of CANDECOMP/PARAFAC and bilinear models with constant modulus constraints," *IEEE Trans. Signal Process.*, vol. 52, no. 9, pp. 2625–2636, Sep. 2004.
- [40] L. De Lathauwer, "A link between the canonical decomposition in multilinear algebra and simultaneous matrix diagonalization," *SIAM J. Matrix Anal. Appl.*, vol. 28, no. 3, pp. 642–666, 2006.
- [41] X. Liu and N. D. Sidiropoulos, "Cramér-Rao lower bounds for low-rank decompositions of multidimensional arrays," *IEEE Trans. Signal Process.*, vol. 49, no. 9, pp. 2074–2089, Sep. 2001.
- [42] T.-J. Shan, M. Wax, and T. Kailath, "On spatial smoothing for direction-of-arrival estimation of coherent signals," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 33, no. 4, pp. 806–811, Aug. 1985.
- [43] N. D. Sidiropoulos and X. Liu, "Identifiability results for blind beamforming in incoherent multipath with small delay spread," *IEEE Trans. Signal Process.*, vol. 49, no. 1, pp. 228–236, Jan. 2001.

SNR Estimation in a Non-Coherent BFSK Receiver With a Carrier Frequency Offset

Syed Ali Hassan and Mary Ann Ingram

Abstract—This correspondence deals with the problem of estimating average signal-to-noise ratio (SNR) for a communication link employing binary frequency shift keying (BFSK) in the presence of a carrier frequency offset (CFO). The transmitted symbols are corrupted by Rayleigh fading and additive white Gaussian noise (AWGN). We treat the CFO as a nuisance parameter and estimate it using a data statistics based estimator. This estimate is then used to design a maximum likelihood (ML) estimator to get the estimates of SNR. We also derive the Cramér-Rao bound (CRB) for the estimators and have shown the performance of both the data-aided and non-data-aided estimators.

Index Terms—BFSK receiver, carrier frequency offset, Rayleigh fading, SNR estimation.

I. INTRODUCTION

Estimates of signal-to-noise ratio (SNR) are used in many wireless receiver functions, including signal detection, power control algorithms, link adaptation, and turbo decoding, etc. Furthermore, if the radios are energy constrained, e.g., if they are in a sensor network, constant envelope modulation and noncoherent demodulation are desirable to reduce circuit consumption of energy. Frequency shift keying (FSK) enables efficient power amplification in the transmitter and a simple receiver design that employs envelope detection. However, the carrier frequency offset (CFO) causes error in the estimation of SNR; this will degrade the performance of systems that depend upon the knowledge of SNR, e.g., amplify and forward (AF) cooperative algorithms [1]. Thus, in this correspondence, we estimate the SNR of a noncoherent BFSK receiver in the presence of a carrier offset, treating the CFO as

a nuisance parameter. The CFO estimation problem is quite tedious to solve because of its highly nonlinear nature, hence analytical methods cannot be directly applied to solve the problem at hand. Therefore, in this correspondence we derive a maximum likelihood estimator for the SNR that uses a moment-based CFO estimator. We also derive the Cramér-Rao lower bound (CRB) for the SNR estimator. We provide two types of SNR estimators: a data-aided (DA) estimator that uses the pilot symbols and a non-data-aided (NDA) estimator that does blind estimation on the received symbols.

Several authors have attacked the problem of estimating the SNR for binary phase shift keying (BPSK) and FSK receivers under perfect synchronization conditions. For example, [2] compares a variety of techniques for SNR estimation in AWGN for M-PSK signals. Many approaches also include the channel effects such as multipath fading and address the issue of SNR estimation for fading channels for BPSK, e.g., in [3]–[6]. In [7], the authors have estimated the average SNR for noncoherent binary FSK (NCBFSK) receiver, assuming a Rayleigh fading channel and unit noise power spectral density. However, in implementations, noise power must also be estimated. The authors in [8] and [9] have derived the SNR estimators for the noncoherent MFSK receivers for Rayleigh as well as block fading channels. However, they assume perfect carrier synchronization at the receiver. The work in [10] and [11] have addressed the problem of calculating the bit error probabilities for NCBFSK systems in the presence of CFO.

The rest of the correspondence is organized as follows. In the next section, we describe the system model and the notations used for the BFSK receiver operating in the presence of CFO. Section III treats the derivations of the SNR estimators for the data-aided scenario in the presence of a Rayleigh fading channel. Section IV considers the estimators for non-data-aided case and in Section V, we will discuss the simulation results for various estimators and overall estimator performance in terms of mean-squared error and the CRB. The correspondence then concludes in Section VI.

II. SYSTEM MODEL

Consider a Rayleigh fading communication system employing binary FSK modulation, where a block of data with k symbols undergoes symbol-rate fading. The received signal observed at the receiver end is given as

$$r(t) = \sqrt{E_s} \alpha(t) \exp(j2\pi(f_c + f_m + \Delta f)t + \theta) + n(t), \quad 0 \leq t \leq T, \quad m \in \{1, 2\} \quad (1)$$

where E_s is the signal power, T is the symbol time, f_c is the carrier frequency, and f_m is the BFSK frequency corresponding to the message signal. The shift in the carrier frequency at the receiver is denoted as Δf and θ is the unknown carrier phase. Without the loss of generality we will set $\theta = 0$, since we are dealing with the noncoherent receiver. The noise at the receiver is $n(t)$, which is AWGN and $\alpha(t)$ is the Rayleigh fading envelope. Thus, the integrator output, matched to the transmitted signal S_m is given as

$$v_m = \int_0^T r(t) \sqrt{E_s} \exp(-j2\pi(f_c + f_m)t) dt \quad (2)$$

where we get the signal part, after simplifying the above equation, as

$$v_{ms} = P \alpha \left[\frac{1 - \exp(-j2\pi\Delta f T)}{j2\pi(m-1) + j2\pi\Delta f T} \right], \quad m \in \{1, 2\} \quad (3)$$

where $P = \frac{E_s T}{2}$. For the sake of simplicity, we assume that the average symbol energy is unity, i.e., $P = 1$, and thus the signal output after the

Manuscript received October 26, 2010; revised January 25, 2011; accepted March 14, 2011. Date of publication March 22, 2011; date of current version June 15, 2011. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Mathini Sellathurai. The authors gratefully acknowledge support for this research from the National Science Foundation under Grant CNS-0721296.

The authors are with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, 30332 USA (e-mail: alihassan@gatech.edu; mai@ece.gatech.edu).

Color versions of one or more of the figures in this correspondence are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TSP.2011.2131137