EURASIP Journal on
Advances in Signal Processing
a SpringerOpen Journal

RESEARCH

Open Access

PARALIND-based identifiability results for parameter estimation via uniform linear array

Xu Liu[1*], Ting Jiang[2], Longxiang Yang[1] and Hongbo Zhu[1]

Abstract

This article applies PARAllel profiles with LINear Dependencies (PARALIND) model to analyze identifiability of parameter estimation in the presence of incoherent multipath via uniform linear array (ULA). New identifiability results are derived based on the uniqueness property of PARALIND model and structure property of ULA. With the strong properties of trilinear model, the proposed identifiability conditions for propagation parameter identification are superior to early studies. We give a new tradeoff between the number of receiving antennae and sampling diversity to ensure parameter identification. Furthermore, a new lower bound of the number of receiving antennae for identifiability is derived. It also shows that the identifiability results is not only determined by traditional factors, such as the number of receiving antennae, oversampling factor or the total number of transmitting paths, but also related to the structure of multipath of sources.

Keywords: Trilinear model, Identifiability, Parameter estimation, Array signal processing

Introduction

Deterministic parameter estimation is one major problem in multi-sensor array system to effectively locate and track various types of signals to minimize interference and maximize intended signal reception, capitalizing various structure property of source signals or (and) received signals [1-4]. The identifiability issue of parameter determination signifies the existence of a unique desired solution under ideal operating conditions and lays the foundation of the capability of estimation techniques. Identifiability results are usually related to the analysis method of the data model and a given algorithm. ESPRIT algorithm, which takes advantage of the rotational invariance property of the uniform linear array (ULA), can be valid to estimate direction of arrivals (DOAs) uniquely only if the number of calibrated receiving antennae is more than the number of sources and all single path signals follow distinct direction to the receiving end [5]. Fonvard/backward spatial smoothing techniques pointed out that $3K/2$ sensor elements should be enough to identify K DOAs of coherent signals [6,7]. In multipath scenario, wireless channel is characterized not only by its DOA but also time delay of the different propagation paths. Van der Veen proposed a joint angle and delay estimation algorithm based on the smoothing method and joint diagonalization technique. A lower bound of number of receiving antennae and oversampling diversity for parameters identification has been presented for the given algorithm [1,8,9]. Recently, Sidiropoulos and Liu [10] linked trilinear decomposition to array signal processing and guaranteed several improved identifiability results of parameter estimation based on PARAFAC analysis, which introduces a new perspective to parameters estimation.

Trilinear data analysis models, such as Tucker3, PARAFAC and PARAllel profiles with LINear Dependencies (PARALIND), were applied into signal processing area in recent years [11-17]. PARALIND model is a kind of trilinear model that was first proposed by Bro et al. [18-20]. It can further be viewed as a new family of PARAFAC models and was developed to extend its usage to problems with linearly dependent factors. Then De Lathauwer and A.L.F. de Almeida introduced the 'Block term decomposition' and 'Constrained block PARAFAC', respectively, which have similar formulations but natural extensions to PARALIND [21,22]. This article links PARALIND analysis to model identifiability of parameter estimation via ULA.

*Correspondence: liuxu@njupt.edu.cn
[1] Jiangsu Key Laboratory of Wireless Communication, College of Telecommunications and Information Engineering, Nanjing University of Posts & Telecommunications, Nanjing, China
Full list of author information is available at the end of the article

Received signals of the ULA, transmitted through incoherent multipath rays of sources with distinct angles and delays, are constructed into PARALIND model. New identifiability results are presented based on the uniqueness issue of PARALIND. The main contributions of this article are listed in the following:

(i) A new 'space-time' tradeoff between the number of receiving antennae and sampling diversity for parameter identification is derived based on the strong uniqueness properties of trilinear model.

(ii) We give a new lower bound of the number of receiving antennae to identify parameters in multipath propagation scenario, which is more superior to early studies.

(iii) Our work shows that the identifiability results for parameters identification are not only determined by some traditional factors, such as the number of receiving antennae, sampling diversity or the number of paths, but also related to the structure of multipath of sources, which was not considered in previous work.

The rest of this article is organized as follows: Section "Data model" lays the data model of array signals in multipath propagation channel. Section "Uniqueness of paralind" gives the basic uniqueness property of PARALIND model. Section "Paralind-based identifiability results for parameter estimation" proposes the main results of parameters identifiability. Some lemmas and theorems will be guaranteed and analyzed. In the last section, we draw the conclusion.

Some notations will be used in this article. $\text{diag}([a, b, \ldots])$ denotes the diagonal matrix with scalar entries $a, b, \ldots$ while $\text{blockdiag}([\mathbf{A}, \mathbf{B}, \ldots])$ denotes the block diagonal matrix with matrix entries $\mathbf{A}, \mathbf{B}, \ldots$. $(\cdot)^T$ and $(\cdot)^\dagger$ stand for transpose and pseudo-inverse, respectively; $\text{vec}(\cdot)$ stacks the columns of its matrix argument in a vector; $\text{unvec}(\cdot)$ is the inverse operation of $\text{vec}(\cdot)$, $\text{unvec}(\mathbf{c}, I, J) = [\mathbf{c}(1 : J), \mathbf{c}(J + 1 : 2J), \ldots, \mathbf{c}((I - 1)J : IJ)]$. $\otimes$ is Kronecker product; $\odot$ denotes the Khatri-Rao product, which is a column-wise Kronecker product. Define $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_R] \in \mathbb{C}^{I \times R}$, $\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_R] \in \mathbb{C}^{J \times R}$, The Khatri-Rao product of $\mathbf{A}$ and $\mathbf{B}$ is:

$$\mathbf{A} \odot \mathbf{B} = \mathbf{a}_1 \otimes \mathbf{b}_1, \ldots, \mathbf{a}_R \otimes \mathbf{b}_R$$
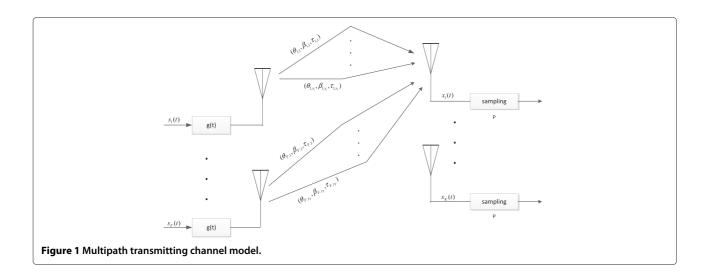
Data model

Figure 1 gives a schematic communication scenario with multipath channel. F sources are transmitting to an array with K antennae through multipath scattering propagation channel. $g(t)$ is the impulse response which collects all temporal aspects, such as pulse shaping, transmitting filter and receiving filter. Signals of fth user follow r_f distinct paths on its way from source to receiver, referred as multipath rays with distinct DOA, transmitting delay and attenuation. The jth path of source f is parameterized by a triple $(\theta_{f,j}, \beta_{f,j}, \tau_{f,j})$, where

$\theta_{f,j}$: DOA
$\beta_{f,j}$: complex path attenuation
$\tau_{f,j}$: transmitting delay

Assume that the ULA is used in the receiving end and the distance d between adjacent elements is equal to (or less than) half of the wavelength of signals. Define r to be the total number of paths of all sources, as $r = \sum_{f=1}^{F} r_f$. Let us conveniently index the multiple rays of sources from 1 to r, starting with all rays associated with the first source and then rays associated with the second source, and so on. Index r parameter triples as $\{(\theta_1, \beta_1, \tau_1), \ldots, (\theta_{r_1}, \beta_{r_1}, \tau_{r_1}), \ldots, (\theta_r, \beta_r, \tau_r)\}$. The array

Figure 1 Multipath transmitting channel model.

manifold matrix $\mathbf{A}_\theta$, time manifold matrix $\mathbf{G}_\tau$ and path attenuation matrix $\mathbf{\Gamma}$ are defined as:

$$\begin{aligned}
\mathbf{A}_\theta &= \left[\mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_{r_1}), \ldots, \mathbf{a}(\theta_r) \right] \in \mathbb{C}^{K \times r} \\
\mathbf{G}_\tau &= \left[\mathbf{g}(\tau_1), \ldots, \mathbf{g}(\tau_{r_1}), \ldots, \mathbf{g}(\tau_r) \right] \in \mathbb{C}^{P \times r} \\
\mathbf{\Gamma} &= \mathrm{diag}(\beta_1, \ldots \beta_{r_1}, \ldots, \beta_r) \in \mathbb{C}^{r \times r}
\end{aligned} \quad (1)$$

where

$$\begin{aligned}
\mathbf{a}(\theta) &= \left[1, e^{j\frac{2\pi d}{\lambda}\sin(\theta)}, \ldots, e^{j\frac{2\pi(K-1)d}{\lambda}\sin(\theta)} \right]^T \\
\mathbf{g}(\tau) &= \left[g\left(\tfrac{1}{P} - \tau\right), g\left(\tfrac{2}{P} - \tau\right), \ldots, g\left(1 - \tfrac{1}{P} - \tau\right) \right]^T
\end{aligned} \quad (2)$$

Received signals can be formulated as follow [1]:

$$\mathbf{X} = (\mathbf{G}_r \odot \mathbf{A}_\theta)\mathbf{\Gamma}(\mathbf{SJ})^T \quad (3)$$

where

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}(0) & \cdots & \mathbf{x}(N-1) \\ \vdots & \vdots & \vdots \\ \mathbf{x}(\tfrac{P-1}{P}) & \cdots & \mathbf{x}(N-1+\tfrac{P-1}{P}) \end{bmatrix} \quad (4)$$

is a $KP \times N$ space-time data matrix collecting samples during N symbol periods with oversampling factor P in the receiving end. $\mathbf{x}$ is a $K \times 1$ array receiving signal. $\mathbf{S}$ is a data matrix of size $N \times F$, collecting N symbols of all users. $\mathbf{J}$ is a selection matrix that joins multipath associated with a given source.

$$\mathbf{J} = \begin{bmatrix} \mathbf{1}_{r_1}^T & 0, \ldots, & 0 \\ & \ddots & \\ 0 & \ldots, & \mathbf{1}_{r_F}^T \end{bmatrix} \quad (5)$$

where $\mathbf{1}_m$ denotes an $m \times 1$ vector with elements 1. Equation (3) is a classical parameterized data model named "incoherent multipath with small delay spread" [1,10]. The propagation parameters, θ_i and τ_i, $i = 1, \ldots, r$, are involved in array manifold matrix $\mathbf{A}_\theta$ and time manifold matrix $\mathbf{G}_\tau$. The multipath structure is indicated by $\mathbf{J}$.

The time delay τ is usually difficult to estimate from $g(t - \tau)$ directly. An alternative approach is to map τ into phase shift ϕ in the frequency domain by discrete Fourier transform (DFT) method [8]. Assume that $g(t)$ is band limited and the sample rate is at or above the Nyquist rate. Take P points DFT of each antenna output over single symbol period. Then the following model is obtained [1]:

$$\bar{\mathbf{X}} = (\mathbf{F}_\phi \odot \mathbf{A}_\theta)\mathbf{\Gamma}(\mathbf{SJ})^T \quad (6)$$

where

$$\mathbf{F}_\phi = \begin{bmatrix} 1 & \ldots & 1 \\ \phi_1 & \ldots & \phi_r \\ \vdots & \vdots & \vdots \\ \phi_1^{P-1} & \ldots & \phi_r^{P-1} \end{bmatrix}, \phi_i = e^{-j2\pi\tau_i/P} \quad (7)$$

The advantage of (7) versus (3) is that, by using DFT method, delays are transformed into certain phase progressions and $\mathbf{G}_\tau$ is converted into a Vandermonde matrix $\mathbf{F}_\phi$, which can provide facility for parameters estimation. Although DFT method may cause some extra error during parameter estimation, van der Veen et al. [8] has informed that this kind of error is very small comparing to the estimation errors that will occurred in the presence of noise.

According to [18], Equation (6) can be viewed as one slice formulation of PARALIND model. The link to PARALIND implies that generic PARALIND model fitting algorithms are directly applicable to deterministic parameter estimation [20]. However, the identifiability of the model pertains to the capability of recovering parameters in the absence of noise. The main work of this article is to investigate new identifiability results for parameter estimation in PARALIND decomposition perspective. Some novel results, such as the tradeoff between the number of receiving antennae and sampling diversity and the lower bound of receiving antennae for parameter identification, are also derived. Firstly, we give the basic uniqueness of PARALIND model.

Uniqueness of PARALIND

The uniqueness of the PARALIND model lays the foundation of its applications. Because of the linear dependence of the loading factors, PARALIND model does not follow directly from the uniqueness property of PARAFAC, but only has partial uniqueness (or essential uniqueness, defined in [23]), which depends on the specifics of the imposed dependency structure along with the adequacy of the factor variation information provided by a given set of data [19]. The uniqueness property of PARALIND was first proposed in [18] and improved by Stegeman and de Almeida [24]. De Lathauwer [23] has given an essential uniqueness theorem more quantitatively. Two new concepts are needed in this theorem.

Definition 1 (*k*-rank) [25]: Consider a matrix $\mathbf{B}$ of size $I \times J$. If every l columns of $\mathbf{B}$ are linearly independent, but this does not hold for every $l + 1$ columns, then the k-rank of $\mathbf{B}$ is l, denoted as $k_\mathbf{B} = l$.

Definition 2 (*k'*-rank) [23]: Assume a partitioned matrix $\mathbf{A} = [\mathbf{A}_1, \ldots, \mathbf{A}_M]$. The *k'*-rank of $\mathbf{A}$, denoted as $\mathrm{rank}_{k'}(\mathbf{A})$

or $k'_{\mathbf{A}}$, is the maximal number r such that any set of r sub-matrices of $\mathbf{A}$ yields a set of linearly independent columns.

Theorem 1 [23]: Rewrite one slice matrix of PARALIND model

$$\mathbf{X} = (\mathbf{A} \odot \mathbf{B})(\mathbf{CH})^T$$

where $\mathbf{A} \in \mathbb{C}^{I \times r}, \mathbf{B} \in \mathbb{C}^{J \times r}, \mathbf{C} \in \mathbb{C}^{K \times F}$. $\mathbf{H}$ is dependence matrix

$$H = \begin{bmatrix} 1,\ldots,1,0,\ldots,0,\ldots,0,\ldots,0 \\ 0,\ldots,0,1,\ldots,1,\ldots,0,\ldots,0 \\ \cdots \\ \underbrace{0,\ldots,0,}_{r_1}\underbrace{0,\ldots,0,}_{r_2}\ldots,\underbrace{1,\ldots,1}_{r_F} \end{bmatrix} = [H_1, H_2, \ldots, H_F] \tag{8}$$

where $\mathbf{H}_f \in \mathbb{C}^{F \times r_f}, f = 1, .., F$ are sub-matrices of $\mathbf{H}$ and $r = \sum_{f=1}^{F} r_f$. $\mathbf{A}, \mathbf{B}$ are partitioned as: $\mathbf{A} = [\mathbf{A}_1, \ldots, \mathbf{A}_F]$, $\mathbf{B} = [\mathbf{B}_1, \ldots, \mathbf{B}_F]$ with the sub-matrices $\mathbf{A}_f \in \mathbb{C}^{I \times r_f}, \mathbf{B}_f \in \mathbb{C}^{J \times r_f}, f = 1, \ldots, F$, compatible with the block structure of $\mathbf{H}$. Suppose that the condition:

$$k'_{\mathbf{A}} + k'_{\mathbf{B}} + k_{\mathbf{C}} \geq 2F + 2 \tag{9}$$

holds and we have an alternative decomposition of $\mathbf{X}$, represented by $(\hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}})$ with $k'_{\hat{\mathbf{A}}}$ and $k'_{\hat{\mathbf{B}}}$ maximal under the given dimensionality constraints. Then there holds $\hat{\mathbf{A}} = \mathbf{A}\boldsymbol{\Pi}_a\boldsymbol{\Delta}_a, \hat{\mathbf{B}} = \mathbf{B}\boldsymbol{\Pi}_b\boldsymbol{\Delta}_b$, where $\boldsymbol{\Pi}_a, \boldsymbol{\Pi}_b$ are block permutation matrices and $\boldsymbol{\Delta}_a, \boldsymbol{\Delta}_b$ are square nonsingular block-diagonal matrices, compatible with the block structure of $\mathbf{A}$ and $\mathbf{B}$.

Theorem 1 presents the uniqueness properties of $\mathbf{A}$ and $\mathbf{B}$. The uniqueness of matrix $\mathbf{C}$ is also studied in [18,21]. Bro et al. [18] gives a demonstration of the uniqueness property of $\mathbf{C}$, provided that $\mathbf{A}, \mathbf{B}$ and $\mathbf{C}$ are full rank. Furthermore, de Lathauwer [21] gives the identifiability result of $\mathbf{C}$ more quantitatively in terms of high-order block tensor decomposition.

Consider Theorem 1 in sub-matrix formulation. Partition $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ to be compatible with the block structure of $\mathbf{A}$ and $\mathbf{B}$, as: $\hat{\mathbf{A}} = [\hat{\mathbf{A}}_1, \ldots, \hat{\mathbf{A}}_F], \hat{\mathbf{B}} = [\hat{\mathbf{B}}_1, \ldots, \hat{\mathbf{B}}_F]$. According to Theorem 1, it directly follows:

$$\begin{aligned}[\hat{\mathbf{A}}_1, \ldots, \hat{\mathbf{A}}_F] &= [\mathbf{A}_1, \ldots, \mathbf{A}_F]\,\boldsymbol{\Pi}_a\boldsymbol{\Delta}_a \\ &= [\mathbf{A}_1, \ldots, \mathbf{A}_F]\,\text{blockdiag}(\mathbf{U}_1, \ldots, \mathbf{U}_F)\end{aligned} \tag{10}$$

$$\begin{aligned}[\hat{\mathbf{B}}_1, \ldots, \hat{\mathbf{B}}_F] &= [\mathbf{B}_1, \ldots, \mathbf{B}_F]\,\boldsymbol{\Pi}_b\boldsymbol{\Delta}_b \\ &= [\mathbf{B}_1, \ldots, \mathbf{B}_F]\,\text{blockdiag}(\mathbf{V}_1, \ldots, \mathbf{V}_F)\end{aligned}$$

where $\mathbf{U}_f \in \mathbb{C}^{r_f \times r_f}, \mathbf{V}_f \in \mathbb{C}^{r_f \times r_f}, f = 1, \ldots, F$ are $2F$ nonsingular square matrices. It follows

$$\begin{aligned}\hat{\mathbf{A}}_m &= \mathbf{A}_m\mathbf{U}_m \\ \hat{\mathbf{B}}_m &= \mathbf{B}_m\mathbf{V}_m, \quad m = 1, \ldots F\end{aligned} \tag{11}$$

Equation (11) gives another representation for uniqueness property of PARALIND model. It shows that the column space of $\mathbf{A}_m$ and $\mathbf{B}_m$ are unique. However, it also implies when condition (9) is satisfied, mode matrices $\mathbf{A}$ and $\mathbf{B}$ are suffered from rotation ambiguity, characterized by $\mathbf{U}_f$ and $\mathbf{V}_f$. Bro et al. [20] has pointed out that PARALIND model can also give uniqueness results if some of its mode matrices have theoretically motivated structural constraints. Due to the structure property of multi-sensor array, we give the PARALIND-based identifiability results for parameter estimation in the next section.

PARALIND-based identifiability results for parameter estimation

As we mentioned, data model (6) can be linked to PARALIND analysis. Array manifold $\mathbf{A}_\theta$, time manifold $\mathbf{F}_\phi$, data matrix $\mathbf{S}$ and selection matrix $\mathbf{J}$ play the roles of $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{H}$ in Theorem 1, respectively. Since the attenuation matrix $\boldsymbol{\Gamma}$ only leads to column scaling of $\mathbf{A}_\theta$ and $\mathbf{F}_\phi$, which will not affect the identifiability results. Therefore, we simplify $\boldsymbol{\Gamma}$ to be an identity matrix during the following discussion. Then the data model is simplified as:

$$\bar{\mathbf{X}} = (\mathbf{F}_\phi \odot \mathbf{A}_\theta)(\mathbf{SJ})^{\mathrm{T}} \tag{12}$$

With the structure of $\mathbf{A}_\theta$ and $\mathbf{F}_\phi$, if these two matrices are uniquely determined, parameters θ and τ are determined naturally. According to Theorem 1, k-rank and k'-rank play important roles in the uniqueness issue of PARALIND. Firstly, we present two lemmas to determine the k-rank and k'-rank of a Vandermonde matrix.

Lemma 1 (k-rank of Vandermonde matrix) [26]

Consider an $I \times r$ Vandermonde matrix $\mathbf{A}$ with distinct nonzero generators $\alpha_1, \alpha_2, \ldots, \alpha_r \in \mathbb{C}$

$$\mathbf{A} = \begin{bmatrix} 1, & \ldots, 1 \\ \alpha_1, & \ldots, \alpha_r \\ \vdots & \\ \alpha_1^{I-1}, & \ldots, \alpha_r^{I-1} \end{bmatrix} \tag{13}$$

$\mathbf{A}$ is not only full rank but also full k-rank, and $k_A = r_A = \min(I, r)$.

Lemma 2 (k'-rank of Vandermonde matrix)

Consider the Vandermonde matrix $\mathbf{A}$ in (13). Partition $\mathbf{A}$ to F sub-matrices, as $\mathbf{A} = [\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_F]$, where $\mathbf{A}_f$ is

of size $I \times r_f$ and $r = r_1 + r_2 + \cdots + r_F$. Resort $r_1, \ldots, r_F$ in descent order and assume that $r_1 > r_2 > \cdots > r_F$. The k'-rank of **A** can be determined as:

$$k'_A = K, \quad \text{where} \sum_{f=1}^{K} r_f \leq I < \sum_{f=1}^{K+1} r_f \quad (14)$$

Note that the k'-rank of **A** is determined not only by its dimension but also the partition structure. If $r_1 = r_2 =, \cdots, = r_F = 1, k_A = k'_A$.

Proof. See Appendix. □

PARALIND-based identifiability result

According to (6), both $\mathbf{A}_\theta$ and $\mathbf{F}_\phi$ are Vandermonde matrices. Capitalizing on the property of PARALIND model and Vandermonde structure, we have the following theorem.

Theorem 2:

Consider data model (12)

$$\bar{\mathbf{X}} = (\mathbf{F}_\phi \odot \mathbf{A}_\theta)(\mathbf{SJ})^{\mathrm{T}}$$

Partition $\mathbf{F}_\phi$ and $\mathbf{A}_\theta$ to F sub-matrices compatible with the structure of **J**, as: $\mathbf{F}_\phi = \left[\mathbf{F}_\phi^1, \ldots, \mathbf{F}_\phi^F \right]$ and $\mathbf{A}_\theta = \left[\mathbf{A}_\theta^1, \ldots, \mathbf{A}_\theta^F \right]$, where $\mathbf{F}_\phi^f$ is $P \times r_f$ and $\mathbf{A}_\theta^f$ is $K \times r_f$. Suppose that the condition

$$k'_{\mathbf{A}_\theta} + k'_{\mathbf{F}_\phi} + k_\mathbf{S} \geq 2F + 2 \quad (15)$$

holds. Then $\mathbf{F}_\phi$ and $\mathbf{A}_\theta$ can be uniquely determined from $\bar{\mathbf{X}}$. The related parameters, DOA θ and delay spread τ, are identifiable.

Proof. See Appendix. □

Although condition (15) in Theorem 2 and condition (9) in Theorem 1 are identical, the identifiability results of these two theorems are different. Theorem 1 shows that $\mathbf{A}_\theta$ and $\mathbf{F}_\phi$ only have "column-space" uniqueness due to the rotation ambiguity in their sub-matrices when condition (9) is satisfied. However, Theorem 2 indicates that Vandermonde matrices $\mathbf{A}_\theta$ and $\mathbf{F}_\phi$ can be uniquely determined from $\bar{\mathbf{X}}$ (no rotation ambiguity) under condition (15). According to the structure of array-manifold matrix $\mathbf{A}_\theta$ and time-manifold $\mathbf{F}_\phi$, the elements of the first row in these two matrices are equal to 1. The scaling ambiguous of the estimated matrices can be removed by normalizing the elements of $\mathbf{A}_\theta$ and $\mathbf{F}_\phi$ with respect to elements of the first row during parameter estimation.

Remark 1: As a special case of (15), we assume that data matrix **S** is full k-rank, as $k_\mathbf{S} = F$. It is achievable when the receiving antennae collect enough symbols

for parameter estimation. Then condition (15) becomes $k'_{\mathbf{A}_\theta} + k'_{\mathbf{F}_\phi} \geq F + 2$. According to the definition of k'-rank, the maximum k'-rank of $\mathbf{A}_\theta$ or $\mathbf{F}_\phi$ is F. Then it requires that $\min(k'_{\mathbf{A}_\theta}, k'_{\mathbf{F}_\phi}) \geq 2$. This lower bound is similar to the identifiability requirement of PARAFAC model, which uses k-rank instead of k'-rank (see Ref. [11]). Furthermore, according to Lemma 2, the k'-rank of $\mathbf{A}_\theta$ and $\mathbf{F}_\phi$ can be represented by $r_1, \ldots, r_F$. Then the minimum value of P and K can be determined as:

$$\min(P, K) \geq \sum_{i=1}^{2} r_i \quad (16)$$

where $r_1, \ldots, r_F$ are descent sorted. Condition (16) shows an interesting result that, to ensure identifiability of parameters estimation, the minimum number of receiving antennae K and oversampling factor P is related to not only the number of sources, but also the number of paths of "some" sources.

Remark 2: Condition (15) shows that the identifiability of data model is determined by $k'_{\mathbf{A}_\theta}$ and $k'_{\mathbf{F}_\phi}$. Lemma 2 also indicates that the k'-rank of $\mathbf{A}_\theta$ and $\mathbf{F}_\phi$ are related to the structure of their sub-matrices, which is compatible with the multipath structure of sources. Therefore, the identifiability result based on PARALIND analysis are not only determined by traditional factors, such as P, K and r, but also related to the structure of multipath, denoted as **J**. The following example can show this phenomenon:

Assume that the number of receiving antennae $K = 4$ and the oversampling factor $P = 6$. Six paths from three sources are arriving at the receiving end. Consider the following two cases:

(1) The number of paths of each source is: $r_1 = 2, r_2 = 2, r_3 = 2$. In this case, $k'_{\mathbf{A}_\theta} = 2$ and $k'_{\mathbf{F}_\phi} = 3$. It has

$$k'_{\mathbf{A}_\theta} + k'_{\mathbf{F}_\phi} = 5 \geq F + 2 = 5$$

According Theorem 2, parameter identification is achievable.

(2) The number of paths of each source is: $r_1 = 3, r_2 = 2, r_3 = 1$. In this case, $k'_{\mathbf{A}_\theta} = 1$ and $k'_{\mathbf{F}_\phi} = 3$. It has

$$k'_{\mathbf{A}_\theta} + k'_{\mathbf{F}_\phi} = 4 < F + 2 = 5$$

Theorem 2 is violated.

It shows that, although the receiving antennae number, oversampling factor and the multipath remain the same, the identifiability results may be different due to the multipath structure. Furthermore, [20] has shown that the dependence matrix **J** can be uniquely obtained in the

PARALIND model by trilinear decomposition method. Therefore, the receive array can get the information of multipath directly from the data model.

PARALIND-based identifiability result with smoothing technique

The identifiability result of (15) can be alleviated by introducing spatial and temporal smoothing techniques from taking advantage of Vandermonde structure in array manifold matrix $\mathbf{A}_\theta$ and time manifold matrix $\mathbf{F}_\phi$. Take 'temporal smoothing' for example. Rewrite (15)

$$\bar{\mathbf{X}} = (\mathbf{F}_\phi \odot \mathbf{A}_\theta)(\mathbf{SJ})^{\mathrm{T}}$$

Construct L matrices of size $MK \times N$

$$
\begin{aligned}
\bar{\mathbf{X}}^{(l)} &= \bar{\mathbf{X}}((l-1)K+1:(M+l-1)K,:) \\
&= \begin{bmatrix}
\bar{\mathbf{X}}((l-1)K+1:lK,:) \\
\bar{\mathbf{X}}(lK+1:(l+1)K,:) \\
\bar{\mathbf{X}}((l+1)K+1:(l+2)K,:) \\
\vdots \\
\bar{\mathbf{X}}((l+M-2)K+1:(l+M-1)K,:)
\end{bmatrix} \\
&= \begin{bmatrix}
\mathbf{F}_\phi(l,:) \odot \mathbf{A}_\theta \\
\mathbf{F}_\phi(l+1,:) \odot \mathbf{A}_\theta \\
\mathbf{F}_\phi(l+2,:) \odot \mathbf{A}_\theta \\
\vdots \\
\mathbf{F}_\phi(l+M-1,:) \odot \mathbf{A}_\theta
\end{bmatrix} (\mathbf{SJ})^{\mathrm{T}} \\
&= (\mathbf{F}_\phi(l:l+M-1,:) \odot \mathbf{A}_\theta)(\mathbf{SJ})^{\mathrm{T}}, \quad l=1,\ldots,L
\end{aligned}
\tag{17}
$$

where $\mathbf{A}(a:b,:)$ stands for rows a to b (inclusive) of $\mathbf{A}$. L is defined as smoothing factor and $M = P - L + 1$. Due to the Vandermonde structure, it holds that

$$
\begin{aligned}
\mathbf{F}_\phi(l:l+M-1,:) &= \begin{bmatrix}
\phi_1^{l-1} & \cdots & \phi_r^{l-1} \\
\vdots & \vdots & \vdots \\
\phi_1^{l+M-2} & \cdots & \phi_r^{l+M-2}
\end{bmatrix} \\
&= \begin{bmatrix}
1 & \cdots & 1 \\
\phi_1 & \cdots & \phi_r \\
\vdots & \vdots & \vdots \\
\phi_1^{M-1} & \cdots & \phi_r^{M-1}
\end{bmatrix} \mathrm{diag}([\phi_1^{l-1}, \phi_2^{l-1}, \ldots, \phi_r^{l-1}]) \\
&= \mathbf{F}_\phi^M \mathrm{diag}(\boldsymbol{\phi}^{l-1})
\end{aligned}
\tag{18}
$$

where $\boldsymbol{\phi}^{l-1}$ denotes $[\phi_1^{l-1}, \phi_2^{l-1}, \ldots, \phi_r^{l-1}]$ and $\mathbf{F}_\phi^M = \mathbf{F}_\phi(1:M,:)$, Substitute (18) into (17)

$$
\begin{aligned}
\bar{\mathbf{X}}^{(l)} &= \left(\left(\mathbf{F}_\phi^M \mathrm{diag}(\boldsymbol{\phi}^{l-1})\right) \odot \mathbf{A}_\theta\right)(\mathbf{SJ})^{\mathrm{T}} \\
&= \left(\left(\mathbf{F}_\phi^M \odot \mathbf{A}_\theta\right) \mathrm{diag}(\boldsymbol{\phi}^{l-1})\right)(\mathbf{SJ})^{\mathrm{T}}
\end{aligned}
\tag{19}
$$

Lay out L matrices $\bar{\mathbf{X}}^{(l)}, l = 1, \ldots, L$ vertically and construct a new matrix $\tilde{\mathbf{X}}$ of size $LMK \times N$

$$
\begin{aligned}
\tilde{\mathbf{X}} &= \begin{bmatrix}
\bar{\mathbf{X}}^{(1)} \\
\vdots \\
\bar{\mathbf{X}}^{(L)}
\end{bmatrix} = \begin{bmatrix}
\left(\left(\mathbf{F}_\phi^M \odot \mathbf{A}_\theta\right) \mathrm{diag}(\boldsymbol{\phi}^0)\right) \\
\vdots \\
\left(\left(\mathbf{F}_\phi^M \odot \mathbf{A}_\theta\right) \mathrm{diag}(\boldsymbol{\phi}^{L-1})\right)
\end{bmatrix} (\mathbf{SJ})^{\mathrm{T}} \\
&= \left(\mathbf{F}_\phi^L \odot \left(\mathbf{F}_\phi^M \odot \mathbf{A}_\theta\right)\right)(\mathbf{SJ})^{\mathrm{T}}
\end{aligned}
\tag{20}
$$

where $\mathbf{F}_\phi^L = \mathbf{F}_\phi(1:L,:)$. Define $\mathbf{A}_\theta^M = \mathbf{F}_\phi^M \odot \mathbf{A}_\theta$. It follows that

$$\tilde{\mathbf{X}} = \left(\mathbf{F}_\phi^{\mathbf{L}} \odot \mathbf{A}_\theta^{\mathbf{M}}\right)(\mathbf{SJ})^{\mathrm{T}} \tag{21}$$

The smoothing data $\tilde{\mathbf{X}}$ can also be modeled as PAR-ALIND. The main difference between (21) and (15) is that model matrices $\mathbf{F}_\phi$ and $\mathbf{A}_\theta$ in (15) is replaced by $\mathbf{F}_\phi^L$ and $\mathbf{A}_\theta^M$. Parameters can also be determined if $\mathbf{F}_\phi^L$ and $\mathbf{A}_\theta^M$ are uniquely decomposed from $\tilde{\mathbf{X}}$. Before discussing the identifiability result of this smoothed model, we need the following lemma:

Lemma 3 (*k*'-rank of Khatri-Rao product of Vandermonde matrix)

Consider two Vandermonde matrices $\mathbf{A} \in \mathbb{C}^{I \times r}$ and $\mathbf{B} \in \mathbb{C}^{J \times r}$ with distinct nonzero generators. $\mathbf{A}$ and $\mathbf{B}$ are partitioned as $\mathbf{A} = [\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_F]$ and $\mathbf{B} = [\mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_F]$, where $\mathbf{A}_f$ is of size $I \times r_f$ and $\mathbf{B}_f$ is of size $J \times r_f$, respectively, and $r = r_1 + r_2 + \cdots + r_F$. Resort $r_1, \ldots, r_F$ in descent order, as $r_1 > r_2 > \cdots > r_F$. If

$$I + J \geq \sum_{f=1}^{K} r_f + 1 \tag{22}$$

then $k'_{\mathbf{A} \odot \mathbf{B}} \geq K$.

Proof. See Appendix ☐

Theorem 3: Consider the smoothed data model (21)

$$\tilde{\mathbf{X}} = \left(\mathbf{F}_\phi^L \odot \mathbf{A}_\theta^M\right)(\mathbf{SJ})^{\mathrm{T}}$$

where $\mathbf{A}_\theta^M = \mathbf{F}_\phi^M \odot \mathbf{A}_\theta$. $\mathbf{F}_\phi^L$ and $\mathbf{F}_\phi^M$, presented in (31) and (28), are of size $L \times \sum_{f=1}^{F} r_f$ and $M \times \sum_{f=1}^{F} r_f$, respectively. Assume that L is selected as $L = \sum_{f=1}^{R} r_f, R \in [2, F]$ and $\mathbf{S}$ is full k-rank. Suppose that the conditions

$$
\begin{cases}
P + K \geq \sum_{i=1}^{R} r_i + \sum_{i=1}^{F+2-R} r_i, & R \in [2, F] \\
\min(K, P) \geq 2
\end{cases}
\tag{23}
$$

hold. Then $\mathbf{F}_\phi^L$, $\mathbf{F}_\phi^M$ and $\mathbf{A}_\theta$ can be uniquely determined from $\tilde{\mathbf{X}}$. Parameters are identifiable.

Proof. See Appendix □

Remark 3: Theorem 3 is guaranteed by smoothing matrix $\mathbf{F}_\phi$ based on its Vandermonde structure. Note that array manifold $\mathbf{A}_\theta$ is also a Vandermonde matrix. Duality simplifies to symmetry, similar formulation can be obtained by smoothing $\mathbf{A}_\theta$, known as 'spatial smoothing':

$$\hat{\mathbf{X}} = \left(\mathbf{A}_\theta^L \odot \mathbf{F}_\phi^M\right)(\mathbf{SJ})^{\mathrm{T}} \tag{24}$$

where $\mathbf{F}_\phi^M = \mathbf{F}_\phi \odot \mathbf{A}_\theta^M$, $\mathbf{A}_\theta^M = \mathbf{A}_\theta(1:M,:)$, and $\mathbf{A}_\theta^L = \mathbf{A}_\theta(1:L,:)$. Note that data model (24) has the same formulation as (21). It implies that the identifiability results of Theorem 3 is also available when $\mathbf{A}_\theta$ is smoothed instead, while R is the k'-rank of $\mathbf{A}_\theta^L$.

Remark 4: Condition (23) gives a new tradeoff between the number of sensors K and oversampling factor P, referred as "space-time" tradeoff, to achieve parameter identifiability. As a special case of (23), two antennae are sufficient for r path when the oversampling factor P is more than $\sum_{i=1}^{R} r_i + \sum_{i=1}^{F+2-R} r_i - 2$. In Theorem 3, The lower bound of choice of receiving antennae K and sampling diversity P is much superior to that in Theorem 2, discussed in Remark 1. It implies that smoothing technique can further improve the identifiability results of data model. It also shows that the system is capable of supporting many more paths than sensors, provided enough sampling diversity. As the complete symmetry in the roles of P and K, limited samples are also available for r path when enough antennae are used in the receiving end.

Remark 5: Rewrite condition (23)

$$P + K \geq \sum_{i=1}^{R} r_i + \sum_{i=1}^{F+2-R} r_i, \quad R \in [2, F] \tag{25}$$

Similar to Remark 2, the value of P plus K is also related to the structure of multipath, denoted as $r_1, \ldots, r_F$. Moreover, it is of interest that the lower bound of P plus K is varied along R, the k'-rank of $\mathbf{F}_\phi^L$ (or $\mathbf{A}_\theta^L$ in (24)). Now, we prove that the minimum lower bound of (25) can be achieved in the condition of $R = 2$ or $R = F$. Define a function of variable R: $f(R) = \sum_{i=1}^{R} r_i + \sum_{i=1}^{F+2-R} r_i$. Let $\Delta f(R) = f(R) - f(2), R \in (2, F]$. Then we wish to prove that

$$\begin{cases} \Delta f(R) \geq 0, & R \in (2, F) \\ f(F) = f(2) \end{cases}$$

It follows

$$\begin{aligned} \Delta f(R) &= \sum_{i=1}^{R} r_i + \sum_{i=1}^{F+2-R} r_i - \sum_{i=1}^{2} r_i - \sum_{i=1}^{F} r_i \\ &= \left(\sum_{i=1}^{R} r_i - \sum_{i=1}^{2} r_i\right) - \left(\sum_{i=1}^{F} r_i - \sum_{i=1}^{F+2-R} r_i\right) \\ &= \sum_{i=3}^{R} r_i - \sum_{i=F-R+3}^{F} r_i \\ &= \sum_{i=1}^{R-2} (r_{R-i+1} - r_{F-i+1}) \end{aligned}$$

Note that $r_1 \geq r_2 \geq, \cdots, \geq r_F$ and $R \leq F$. It is clear that $r_{R-i+1} \geq r_{F-i+1}$, Therefore, $\Delta f(R) \geq 0$. Since $\Delta f(F) = \sum_{i=3}^{F} r_i - \sum_{i=3}^{F} r_i = 0$, we have $f(F) = f(2)$. This result gives a relationship between the smoothing factor R and the minimum value of P plus K in parameter estimation when r path is considered. Since the cost of parameter estimation is usually related to the number of receiving antenna and the oversampling factor, the result also implies that the cost of parameter estimation can be decreased when the smoothing factor is properly selected.
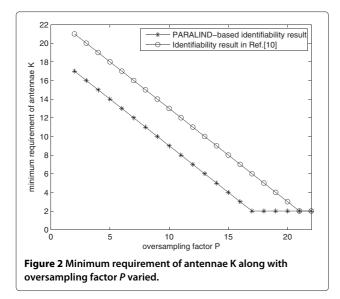
Remark 6: If the multiplication $\mathbf{SJ}$ is considered as a whole matrix, the data model (12) can be modeled as PARAFAC. However, since matrix multiplication $\mathbf{SJ}$ has collinear columns due to the structure of $\mathbf{J}$. According to the uniqueness property of PARAFAC [25], uniqueness of the given model cannot be guaranteed so that meaningful results of parameter identifiability may not be derived directly based on PARAFAC model. Sidiropoulos and Liu [10] utilizes smoothing technique to improve the k-rank of $\mathbf{SJ}$ and gives the identifiability results of (12) based on PARAFAC model. Here we will show that condition (23) is superior to that in [10]. Define matrix $\mathbf{C} = \mathbf{SJ}$. Sidiropoulos and Liu [10] presented that the model (12) is identifiable, provided that

$$K + P + k_{\mathbf{C}} \geq 2r + 2 \tag{26}$$

Note that $\mathbf{C}$ has collinear columns. According to the definition of k-rank, the k-rank of $\mathbf{C}$ is equal to 1. Condition (26) becomes $K + P \geq 2r + 1$. Because $r = \sum_{i=1}^{F} r_i$. We have

$$2r + 1 > \sum_{i=1}^{R} r_i + \sum_{i=1}^{F+2-R} r_i$$

It can be concluded that condition (23) is more relaxed than condition (26). The following example can guarantee the above improvement. Assume that four users are in consideration. The multipath of users are $r_1 = 5, r_2 = 3$, $r_3 = 2, r_4 = 1$, respectively. Figure 2 depicts the minimum requirement of receiving antennae K in these two conditions along with oversampling factor P varied.

Figure 2 Minimum requirement of antennae K along with oversampling factor *P* varied.

Conclusion

This article has discussed identifiability issue of deterministic parameters estimation via multi-sensor array based on trilinear decomposition theory. With the uniqueness property of PARALIND model, new identifiability results are guaranteed, which are more superior to early studies. According to the proposed identifiability conditions, a new "space-time" tradeoff between the number of receiving antennae and sampling diversity for parameters identification is presented, and it shows that even two receiving antennae are sufficient for identifying parameters of r path, provided sufficient sampling diversity available. Besides, we find that the identifiability conditions is not only determined by some traditional factors, such as the number of receiving antennae, oversampling factors or number of paths, but also related to the multipath structure of each source, which was not considered in previous work.

Appendix
Proof of Lemma 2

According to the definition of k'-rank, if $k'_A = K$, it means that any K sub-matrices of **A** yield a set of linearly independent columns but it cannot support $K + 1$ sub-matrices. Let $\tilde{\mathbf{A}}$ be an $I \times \sum_{i=1}^{K} \tilde{r}_i$ matrix including any K sub-matrices of **A**, as $\tilde{\mathbf{A}} = \left[\tilde{\mathbf{A}}_1, \ldots, \tilde{\mathbf{A}}_K \right]$, where $\tilde{\mathbf{A}}_1, \ldots, \tilde{\mathbf{A}}_K$ are randomly selected from $\mathbf{A}_1, \ldots, \mathbf{A}_F$ and $\tilde{\mathbf{A}}_i \neq \tilde{\mathbf{A}}_j, i, j \in [1, F], i \neq j$. Note that $\tilde{\mathbf{A}}_k$ is of size $I \times \tilde{r}_k$ with Vandermonde structure and in most cases we have $\tilde{\mathbf{A}}_k \neq \mathbf{A}_k$ and $\tilde{r}_k \neq r_k$. With the assumption of $r_1 > r_2 > \cdots > r_F$, it follows that $\sum_{i=1}^{K} \tilde{r}_i \leqslant \sum_{i=1}^{K} r_i$. Since $I \geqslant \sum_{f=1}^{K} r_f$ guarantees $I \geqslant \sum_{f=1}^{K} \tilde{r}_f$. According to Lemma 1, $\tilde{\mathbf{A}}$ is full column rank. It implies that any K

sub-matrices of **A** are guaranteed to yield a set of linearly independent columns so that $k'_A \geqslant K$. On the other hand, define $\hat{\mathbf{A}}$ as $\hat{\mathbf{A}} = [\mathbf{A}_1, \ldots, \mathbf{A}_K, \mathbf{A}_{K+1}]$. $\hat{\mathbf{A}}$ is a $I \times \sum_{i=1}^{K+1} r_i$ Vandermonde matrix with $K + 1$ sub-matrices of **A** and, $k_{\hat{\mathbf{A}}} = \min \left(I, \sum_{i=1}^{K+1} r_i \right) = I$. It means that we can find $K + 1$ sub-matrices of **A** which yields a set of dependent columns, so that $k'_A < K + 1$. Therefore, $k'_A = K$. The proof is complete. ∎

Proof of Theorem 2

Before proving Theorem 2, we need the following Lemma:

Lemma 4 [27] Consider a matrix decomposition $\mathbf{X} = \mathbf{A}\mathbf{B}^{\mathbf{T}}$, where $\mathbf{A} \in \mathbb{C}^{I \times F}$ is a Vandermonde matrix with distinct nonzero generator and $\mathbf{B} \in \mathbb{C}^{J \times F}$ is a 'tall' or 'square' matrix with full column rank. Suppose that the condition $I \geqslant F + 1$ holds and then **A** and **B** can be uniquely decomposed from **X** under permutation and scaling ambiguous. It means that any other alternative decomposition of **X**, denoted as $\mathbf{X} = \bar{\mathbf{A}}\bar{\mathbf{B}}^T$ in which $\bar{\mathbf{A}} \in \mathbb{C}^{I \times F}$ has Vandermonde strucure and $\bar{\mathbf{B}} \in \mathbb{C}^{J \times F}$ is full column rank, is related to **A** and **B** via $\bar{\mathbf{A}} = \mathbf{A}\mathbf{\Pi}_A\mathbf{\Delta}_A, \bar{\mathbf{B}} = \mathbf{B}\mathbf{\Pi}_B\mathbf{\Delta}_B$, where $\mathbf{\Pi}_A, \mathbf{\Pi}_B$ are permutation matrices and $\mathbf{\Delta}_A, \mathbf{\Delta}_B$ are diagonal scaling matrices with nonzero elements.

According to Theorem 1, when condition (15) is followed, we have $\hat{\mathbf{F}}_\phi^f = \mathbf{F}_\phi^f \mathbf{U}_f, \hat{\mathbf{A}}_\theta^f = \mathbf{A}_\theta^f \mathbf{V}_f, f = 1, \ldots F$, where $\mathbf{U}_1, \mathbf{V}_1, \ldots, \mathbf{U}_F, \mathbf{V}_F$ are $2F$ nonsingular square matrices. Note that any subset of columns of a Vandermonde matrix forms a Vandermonde matrix. Therefore, $\mathbf{F}_\phi^f, \mathbf{A}_\theta^f, f = 1, \ldots, F$ are all with Vandermonde structure. Lemma 4 provides that $\mathbf{F}_\phi^1, \mathbf{A}_\theta^1, \ldots, \mathbf{F}_\phi^F, \mathbf{A}_\theta^F$ can be uniquely determined from $\hat{\mathbf{F}}_\phi^1, \hat{\mathbf{A}}_\theta^1, \ldots, \hat{\mathbf{F}}_\phi^F, \hat{\mathbf{A}}_\theta^F$ only if the following conditions are satisfied:

$$\begin{cases} P \geqslant r_f + 1 \\ K \geqslant r_f + 1 \end{cases} \quad f = 1, \ldots, F \quad (27)$$

Assume that $r_1 > r_2 > \cdots > r_F$. Then conditions (27) becomes

$$\begin{cases} P \geqslant r_1 + 1 \\ K \geqslant r_1 + 1 \end{cases} \quad (28)$$

Remark 1 derives that the minimum of P and K should be larger than $r_1 + r_2$. Since r_2 is no less than 1(multiple source assumption), it means that $\min(P, K) \geqslant r_1 + 1$ so that (15) is a sufficient condition for (28). Therefore, $\mathbf{F}_\phi^f, \mathbf{A}_\theta^f$ can be uniquely determined from $\hat{\mathbf{F}}_\phi^f, \hat{\mathbf{A}}_\theta^f$ when condition (15) is satisfied. Then $\mathbf{F}_\phi, \mathbf{A}_\theta$ can be uniquely obtained from $\mathbf{F}_\phi^f, \mathbf{A}_\theta^f$. The proof is complete. ∎

Proof of Lemma 3

We need the following Lemma:

Lemma 5 (full rank of Khatri-Rao product) [28]

Consider $\mathbf{A} \odot \mathbf{B} := [\mathbf{a}_1 \otimes \mathbf{b}_1, \ldots, \mathbf{a}_F \otimes \mathbf{b}_F]$, where $\mathbf{A}$ is of size $I \times F$, $\mathbf{B}$ is of size $J \times F$ and $\mathbf{a}_f, \mathbf{b}_f, f = 1, \ldots, F$ are columns of $\mathbf{A}, \mathbf{B}$. If $r_\mathbf{A} + k_\mathbf{B} \geq F + 1$ or $r_\mathbf{B} + k_\mathbf{A} \geq F + 1$ holds, then $\mathbf{A} \odot \mathbf{B}$ is full column rank, as $r_{\mathbf{A} \odot \mathbf{B}} = F$.

Assume $\mathbf{A}, \mathbf{B}$ are Vandermonde matrices. According to Lemma 1, $r_\mathbf{A} = k_\mathbf{A} = \min(I, F), r_\mathbf{B} = k_\mathbf{B} = \min(J, F)$. As a special case of Lemma 5, the full rank condition of $\mathbf{A} \odot \mathbf{B}$ with Vandermonde assumption is

$$\min(I, F) + \min(J, F) \geq F + 1 \quad (29)$$

Randomly select K sub-matrices $\tilde{\mathbf{C}}_1, \ldots, \tilde{\mathbf{C}}_K$ from $\mathbf{C}$ and construct a new matrix

$$\tilde{C} = \left[\tilde{\mathbf{C}}_1, \ldots, \tilde{\mathbf{C}}_K\right] = \left[\tilde{\mathbf{A}}_1 \odot \tilde{\mathbf{B}}_1, \ldots, \tilde{\mathbf{A}}_K \odot \tilde{\mathbf{B}}_K\right]$$
$$= \left[\tilde{\mathbf{A}}_1, \ldots, \tilde{\mathbf{A}}_K\right] \odot \left[\tilde{\mathbf{B}}_1, \ldots, \tilde{\mathbf{B}}_K\right] = \tilde{\mathbf{A}} \odot \tilde{\mathbf{B}}$$

where $\tilde{\mathbf{A}}_f$ is $I \times \tilde{r}_f$, $\tilde{\mathbf{B}}_f$ is $J \times \tilde{r}_f$, $\tilde{\mathbf{A}}$ is $I \times \sum_{f=1}^K \tilde{r}_f$ and $\tilde{\mathbf{B}}$ is $J \times \sum_{f=1}^K \tilde{r}_f$. Note that here $\tilde{\mathbf{A}}_f \in \{\mathbf{A}_1, \ldots, \mathbf{A}_F\}, \tilde{\mathbf{B}}_f \in \{\mathbf{B}_1, \ldots, \mathbf{B}_F\}$, and $\tilde{\mathbf{A}}_i \neq \tilde{\mathbf{A}}_j, \tilde{\mathbf{B}}_i \neq \tilde{\mathbf{B}}_j, i \neq j$. Similar to the proof procedure of Lemma 2, we only need to show that $\tilde{\mathbf{C}}$ is full column rank under condition (22). It is equivalent to prove:

$$\min\left(I, \sum_{f=1}^K \tilde{r}_f\right) + \min\left(J, \sum_{f=1}^K \tilde{r}_f\right) \geq \sum_{f=1}^K \tilde{r}_f + 1 \quad (30)$$

In the light of (30), consider the following cases

(1) $I \geq \sum_{f=1}^K \tilde{r}_f, J \geq \sum_{f=1}^K \tilde{r}_f$. Then

$$\min\left(I, \sum_{f=1}^K \tilde{r}_f\right) + \min\left(J, \sum_{f=1}^K \tilde{r}_f\right)$$
$$= \sum_{f=1}^K \tilde{r}_f + \sum_{f=1}^K \tilde{r}_f \geq \sum_{f=1}^K \tilde{r}_f + 1$$

Condition (30) is satisfied.

(2) $I < \sum_{f=1}^K \tilde{r}_f, J \geq \sum_{f=1}^K \tilde{r}_f$ or $I \geq \sum_{f=1}^K \tilde{r}_f, J < \sum_{f=1}^K \tilde{r}_f$. Then

$$\min\left(I, \sum_{f=1}^K \tilde{r}_f\right) + \min\left(J, \sum_{f=1}^K \tilde{r}_f\right)$$
$$= \min(I, J) + \sum_{f=1}^K \tilde{r}_f \geq \sum_{f=1}^K \tilde{r}_f + 1$$

Condition (30) is satisfied.

(3) $I < \sum_{f=1}^K \tilde{r}_f, J < \sum_{f=1}^K \tilde{r}_f$. Then depending on condition (22),

$$\min\left(I, \sum_{f=1}^K \tilde{r}_f\right) + \min\left(J, \sum_{f=1}^K \tilde{r}_f\right) = I + J \geq \sum_{f=1}^K r_f + 1$$

Since $r_1 \geq r_2 \geq, \cdots, \geq r_F$, and $\{\tilde{r}_1, \ldots, \tilde{r}_K\} \subset \{r_1, \ldots, r_F\}$, it holds that $\sum_{f=1}^K r_f + 1 \geq \sum_{f=1}^K \tilde{r}_f + 1$. Condition (30) is satisfied.

Therefore, $\tilde{\mathbf{C}}$ is full column rank, so that $k'_\mathbf{C} \geq K$. The proof is complete. ■

Proof of Theorem 3

Note that $P = L + M - 1$ and $L = \sum_{i=1}^R r_i \leq P$. According to Lemma 2, the k'-rank of $\mathbf{F}_\phi^L$ is R. Condition (23) becomes

$$P + K \geq \sum_{i=1}^R r_i + \sum_{i=1}^{F+2-R} r_i$$
$$\Rightarrow L + M - 1 + K \geq \sum_{i=1}^R r_i + \sum_{i=1}^{F+2-R} r_i$$
$$\Rightarrow \sum_{i=1}^R r_i + M - 1 + K \geq \sum_{i=1}^R r_i + \sum_{i=1}^{F+2-R} r_i \quad (31)$$
$$\Rightarrow K + M \geq \sum_{i=1}^{F+2-R} r_i + 1$$

According to Lemma 3, condition (31) provides that the k'-rank of $\mathbf{A}_\theta^M$ is larger than $F + 2 - R$. Then we have

$$\left.\begin{array}{l} k'_{\mathbf{F}_\phi^L} = R \\ k'_{\mathbf{A}_\theta^M} \geq F + 2 - R \end{array}\right\} \Rightarrow k'_{\mathbf{F}_\phi^L} + k'_{\mathbf{A}_\theta^M} \geq F + 2 \quad (32)$$

With the assumption of $k_\mathbf{S} = F$ and Theorem 1, condition (32) shows that the partial uniqueness of model (21) is achieved. Then we have

$$\hat{\mathbf{F}}_{\phi,f}^L = \mathbf{F}_{\phi,f}^L \mathbf{V}_f^T \quad (33)$$

$$\hat{\mathbf{A}}_{\theta,f}^M = \mathbf{A}_{\theta,f}^M \mathbf{U}_f^T = \left(\mathbf{F}_{\phi,f}^M \odot \mathbf{A}_{\theta,f}\right) \mathbf{U}_f^T \quad (34)$$

where $\mathbf{V}_f, \mathbf{U}_f$ are nonsingular square matrices of size $r_f \times r_f$. Because $L = \sum_{i=1}^R r_i \geq r_f + 1$. According to Lemma 4,

$\mathbf{F}_{\phi,f}^L$ can be uniquely determined from $\hat{\mathbf{F}}_{\phi,f}^L$. Then we wish to prove that $\mathbf{F}_{\phi,f}^M$ and $\mathbf{A}_{\theta,f}$ can be uniquely determined from $\hat{\mathbf{A}}_{\theta,f}^M$. Consider the following cases with different value of r_f

(1) $r_f = 1$

If $r_f = 1$, $\mathbf{F}_{\phi,f}^M$ and $\mathbf{A}_{\theta,f}$ degenerate to vectors $\mathbf{f}_{\phi,f}^M$, $\mathbf{a}_{\theta,f}$, and $\mathbf{U}_f$ degenerates to a scalar u_f. Equation (34) becomes

$$\hat{\mathbf{a}}_{\theta,f}^M = u_f \left(\mathbf{f}_{\phi,f}^M \odot \mathbf{a}_{\theta,f} \right) = u_f \left(\mathbf{f}_{\phi,f}^M \otimes \mathbf{a}_{\theta,f} \right)$$

where $\hat{\mathbf{a}}_{\theta,f}^M \in \mathbb{C}^{MK \times 1}$, $\mathbf{f}_{\phi,f}^M \in \mathbb{C}^{M \times 1}$ and $\mathbf{a}_{\theta,f} \in \mathbb{C}^{K \times 1}$. Because u_f only leads to column scaling of $\mathbf{f}_{\phi,f}^M$ and $\mathbf{a}_{\theta,f}$, which will not affect the identifiability result. We simplify u_f to be 1. Rearrange $\hat{\mathbf{a}}_{\theta,f}^M$ to be a $M \times K$ matrix $\boldsymbol{\Omega}$

$$\boldsymbol{\Omega} = \text{unvec} \left(\hat{\mathbf{a}}_{\theta,f}^M, M, K \right) = \text{unvec} \left(\mathbf{f}_{\phi,f}^M \otimes \mathbf{a}_{\theta,f}, M, K \right)$$
$$= \mathbf{a}_{\theta,f} \left(\mathbf{f}_{\phi,f}^M \right)^T$$

Then $\mathbf{f}_{\phi,f}^M$ and $\mathbf{a}_{\theta,f}$ can be easily determined from $\boldsymbol{\Omega}$ by using singular value decomposition method (SVD) up to scaling ambiguity.

(2) $r_f \geq 2$

It is of interest that Equation (34) is a standard slice matrix formulation of PARAFAC model when $r_f \geq 2$, of which three mode matrices are $\mathbf{F}_{\phi,f}^M$, $\mathbf{A}_{\theta,f}$ and $\mathbf{U}_f$. Recall that the uniqueness condition of PARAFAC model is [11,25,29-31]

$$k_{\mathbf{F}_{\phi,f}^M} + k_{\mathbf{A}_{\theta,f}} + k_{\mathbf{U}_f} \geq 2r_f + 2 \tag{35}$$

$\mathbf{U}_f$ is a $r_f \times r_f$ square nonsingular matrix with full k-rank, as $k_U = r_f$. $\mathbf{F}_{\phi,f}^M \in \mathbb{C}^{M \times r_f}$ and $\mathbf{A}_{\theta,f} \in \mathbb{C}^{K \times r_f}$ are Vandermonde matrices, and their k-ranks can be determined as $k_{\mathbf{F}_{\phi,f}^M} = \min(M, r_f)$, $k_{\mathbf{A}_{\phi,f}} = \min(K, r_f)$. Then condition (35) becomes:

$$\min(M, r_f) + \min(K, r_f) \geq r_f + 2 \tag{36}$$

We now prove that condition (31) is sufficient to (36). Four cases need to be discussed:

(2.1) $M \geq r_f, K \geq r_f$, then
$\min(M, r_f) + \min(K, r_f) = r_f + r_f \geq r_f + 2$. Condition (36) is satisfied.

(2.2) $M < r_f, K \geq r_f$, then
$\min(M, r_f) + \min(K, r_f) = M + r_f$. Condition (36) is satisfied when $M \geq 2$.

However, if $M = 1, L = P$, the structure of $\mathbf{F}_{\phi,f}^M$ shows that model (21) degenerate to $\tilde{\mathbf{X}} = (\mathbf{F}_\phi \odot \mathbf{A}_\theta)(\mathbf{SJ})^T$. According to (31), $K \geq \sum_{i=1}^{F+2-R} r_i$. The k'-rank of $\mathbf{A}_\theta$ is larger than $F + 2 - R$. With the assumption of $L = \sum_{i=1}^R r_i$, the k'-rank of $\mathbf{F}_\phi$ is R. It holds

$$k'_{\mathbf{F}_\phi} + k'_{\mathbf{A}_\theta} + k_S \geq R + F - R + 2 + F = 2F + 2 \tag{37}$$

Theorem 2 shows that $\mathbf{A}_\theta$ and $\mathbf{F}_\phi$ can be uniquely determined from $\tilde{\mathbf{X}}$ under condition (37).

(2.3) $M \geq r_f, K < r_f$, then
$\min(M, r_f) + \min(K, r_f) = r_f + K$. Since $\min(K, P) \geq 2$. $\min(M, r_f) + \min(K, r_f)$ is larger than $r_f + 2$. Condition (36) is satisfied.

(2.4) $M < r_f, K < r_f$, then
$\min(M, r_f) + \min(K, r_f) = M + K \geq \sum_{i=1}^{F+2-R} r_i + 1$. Because $2 \leq R \leq F$, $\sum_{i=1}^{F+2-R} r_i + 1 \geq r_1 + r_2 + 1 \geq r_f + 2$. Condition (36) is satisfied.

Therefore, $\mathbf{F}_{\phi,f}^L$, $\mathbf{F}_{\phi,f}^M$ and $\mathbf{A}_{\theta,f}$ can be uniquely determined from $\hat{\mathbf{F}}_{\phi,f}^L$ and $\hat{\mathbf{A}}_{\theta,f}^M$ under condition (23). This completes the proof. ∎

Competing interests
The authors declare that they have no competing interests.

Author details
[1] Jiangsu Key Laboratory of Wireless Communication, College of Telecommunications and Information Engineering, Nanjing University of Posts & Telecommunications, Nanjing, China. [2] Nanjing Panda Technology CO. LTD, Nanjing, China.

References
1. AJ van der Veen, Algebraic methods for deterministic blind beamforming. Proc. IEEE. **86**, 1987–2008 (1998)
2. H Krim, M Viberg, Two decades of array signal processing research, the parametric approach. IEEE Signal Process. Mag. **13**, 67–94 (1996)
3. ZT Liu, J He, Z Liu, Computationally efficient DOA and polarization estimation of coherent sources with linear electromagnetic vector-sensor array. EURASIP J. Adv. Signal Process (2010). doi:10.1155/2011/490289
4. A Sohl, A Klein, Semiblind channel estimation for IFDMA in case of channels with large delay spreads. Eurasip J. Adv. Signal Process (2010). doi:10.1155/2011/857859

5. R Roy, T Kailath, ESPRIT-Estimation of signal parameters via rotational invariance techniques. IEEE Trans. Acoust. Speech Signal Process. **37**, 984–995 (1989)

6. S Pillai, B Kwon, Forward–backward spatial smoothing techniques for the coherent signal identification. IEEE Trans. Acoust. Speech Signal Process. **37**, 8–15 (1989)

7. DA Linebarger, RD DeGroat, EM Dowling, Efficient direction-finding methods employing forword/backward averaging. IEEE Trans. Signal Process. **42**(8), 2136–2145 (1994)

8. AJ van der Veen, MC van der Veen, A Paulraj, Joint angle and delay estimation using shift-invariance techniques. IEEE Trans. Signal Process. **46**(2), 405–418 (1998)

9. MC van der veen, AJ van der Veen, Estimation of multipath parameters in wireless communications. IEEE Trans. Signal Process. **46**(3), 682–690 (1998)

10. ND Sidiropoulos, X Liu, Identifiability results for blind beamforming in incoherent multipath with small delay spread. IEEE Trans. Signal Process. **49**(1), 228–238 (2001)

11. ND Sidiropoulos, GB Giannakis, R Bro, Blind PARAFAC receivers for DS-CDMA systems. IEEE Trans. Signal Process. **48**(3), 810–823 (2000)

12. ALF de Almeida, G Favier, JCM Mota, Constrained tucker-3 model for blind beamforming. Signal Process. **89**, 1240–1244 (2009)

13. ND Sidiropoulos, GZ Dimie, Blind multiuser detection in W-CDMA systems with large delay spread. IEEE Signal Process. Lett. **8**(3), 87–89 (2001)

14. XF Zhang, DZ Xu, Blind PARAFAC signal detection for polarization sensitive array. Eurasip J. Adv. Signal Process (2007). doi:10.1155/2007/12025

15. X Liu, ZZ Xu, A PARALIND-based blind multiuser detection algorithm in MIMO-CDMA system. J. Syst. Eng. Electron. (China). **33**(2), 404–410 (2011)

16. JL Liang, SY Yang, JY Zhang, 4D near-field source localization using cumulant. Eurasip J. Adv. Signal Process (2007). doi:10.1155/2007/17820

17. LC Carvalho, F Roemer, M Haardt, Multi-dimensional model order selection. Eurasip J. Adv. Signal Process (2011). doi:10.1186/1687-6180-2011-26

18. R Bro, RA Harshman, ND Sidiropoulos, ME Lundy, Modeling multi-way data with linearly dependent loadings, KVL Technical Report, (2005)

19. M Bahram, R Bro, A novel strategy for solving matrix effect in three-way data using parallel profiles with linear dependencies. Anal. Chim. Acta. **584**, 397–402 (2007)

20. R Bro, RA Harshman, ND Sidiropoulos, ME Lundy, Modeling multi-way data with linearly dependent loadings. J. Chemometr. **23**(7-8), 324–340 (2009)

21. L de Lathauwer, Decomposition of a higher-order tensor in block terms-part I: lemmas for partitioned matrices. SIAM J. Matrix Anal. Appl. **30**(3), 1022–1032 (2008)

22. ALF de Almeida, G Favier, JCM Mota, Constrained tensor modeling approach to blind multiple-antenna CDMA schemes. IEEE Trans. Signal Process. **56**(6), 2417–2428 (2008)

23. L de Lathauwer, Decomposition of a higher-order tensor in block terms-part II: definitions and uniqueness. SIAM J. Matrix Anal. Appl. **30**(3), 1033–1066 (2008)

24. A Stegeman, ALF de Almeida, Uniqueness conditions for constrained three-way factor decompositions with linearly dependent loadings. SIAM J. Matrix Anal. Appl. **31**, 1469–1490 (2009)

25. JB Kruskal, Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. Linear Algeb. Appl. **18**, 95–138 (1977)

26. ND Sidiropoulos, GB Giannakis, R Bro, Parallel factor analysis in sensor array processing. IEEE Trans. Signal Process. **48**(8), 2377–2388 (2000)

27. X Liu, ZZ Xu, L Lei, Identification of array signal parameters based on matrix decomposition. J. Appl. Sci. (China). **28**(1), 49–55 (2010)

28. X Guo, D Brie, S Zhu, X Liao, A CANDECOMP/PARAFAC perspective on uniqueness of DOA estimation using a vector sensor array. IEEE Trans. Signal Process. **59**, 3475–3481 (2011)

29. L De Lathauwer, A link between the canonical decomposition in multilinear algebra and simultaneous matrix diagonalization. SIAM J. Matrix Anal. Appl. **28**, 642–666 (2006)

30. A Stegeman, On uniqueness conditions for Candecomp/Parafac and Indscal with full column rank in one mode. Linear Algeb. Appl. **431**, 211–227 (2009)

31. T Jiang, ND Sidiropoulos, Kruskal's permutation lemma and the identification of Candecomp/Parafac and bilinear models with constant modulus constraints. IEEE Trans. Signal Process. **52**, 2625–2636 (2004)