IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 21, 2018, accepted February 22, 2018, date of current version March 28, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2810225

A Quadrilinear Decomposition Method for Direction Estimation in Bistatic MIMO Radar

ZIYI WANG[1], CHANGXIN CAI[1], FANGQING WEN[1,2], (Member, IEEE), AND DONGMEI HUANG[3]

[1]Electronic and Information School, Yangtze University, Jingzhou 434023, China
[2]Key Laboratory of Radar Imaging and Microwave Photonics, Ministry of Education, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China
[3]Department of Training Management, Naval Command College, Nanjing 210016, China

Corresponding author: Changxin Cai (caichangxinjpu@126.com)

This work was supported in part by National Natural Science Foundation, China, under Grant 61701046 and Grant 61471191, and in part by the Electronic and Information School, Yangtze University Innovation Foundation under Grant 2016-DXCX-05.

ABSTRACT We investigate into the problem of joint direction-of-departure (DOD) and direction-of-arrival (DOA) estimation in a multiple-input multiple-output radar, and a novel covariance tensor-based quadrilinear decomposition algorithm is derived in this paper. By taking into account the multidimensional structure of the matched array data, a fourth-order covariance tensor is formulated, which links the problem of joint DOD and DOA estimation to a quadrilinear decomposition model. A quadrilinear alternating least squares (QALSs) technique is applied to estimate the loading matrices, and thereafter automatically paired DODs and DOAs are obtained via the LSs fitting strategy. The proposed QALS algorithm can be regarded as an alternative to the direct parallel factor (PARAFAC) algorithm, which is more flexible than the latter since it can be easily expanded to scenario with spatially colored noise. Moreover, the proposed algorithm has much lower computational complexity than PARAFAC, especially in the presence of large snapshot. Numerical simulations verify the effectiveness of the proposed algorithm.

INDEX TERMS Multiple-input multiple-output radar, direction estimation, tensor decomposition, colored noise.

I. INTRODUCTION

Multiple-input multiple-output (MIMO) radar is an innovative architecture for target detection. MIMO radar system illuminates a given area by emitting mutual waveforms with multiple antennas. By collecting the echoes with multiple antennas, MIMO radar attempts to determine the object parameters in its field of view. A virtual transmitter-object-receiver channel is formed once a matched filter is applied at the receive end. All the individual transmitter-object-receiver channels consist of the virtual aperture, which enables of a MIMO radar to achieve superior performance than the conventional phased radar, e.g., enhanced spatial resolution, improved interference and jamming suppression. Generally, MIMO radar can be divided into two categories in terms of its antenna configurations: statistical MIMO radar and colocated MIMO radar. The former, whose antennas are widely separated in both the transmitting and receiving arrays, can solve the scintillation problem due to the spatial diversity gain. The latter, whose antennas are closely located in the transmitting and receiving arrays, can achieve super-resolution directions estimation due to the coherent processing gain. In this paper, we focus on bistatic MIMO radar, which belong to the latter.

Joint direction-of-departure (DOD) and direction-of-arrival (DOA) estimation is a canonical problem in bistatic MIMO radar. Many efforts have been tried and various matrix-based algorithms have been developed. For instance, MUSIC [1], ESPRIT [2], propagator method [3] and maximum likelihood method [4]. However, the multi-dimensional nature inherit in the matched array data are neglected in these methods. Signal processing tools based on tensor algebra have been turned out to be effective to improve parameter estimation due to their superior noise rejection capability [5], which is known as 'tensor gain'. Typical tensor-based decomposition frameworks include Tucker decomposition and parallel factor (PARAFAC) decomposition. In [6] and [7], Tucker estimators have been derived for joint DOD and DOA estimation in bistatic MIMO radar, which help to achieve an more accurate signal/noise subspace than the traditional singular value decompositon/eigendecompositon methods. PARAFAC estimators have been investigated in [8] and [9], which factorize the tensor

data into sum of rank-one vectors and thus obtain the least squares (LS) estimation of the loading matrices. Compared with the Tucker estimators, PARAFAC estimators are usually computationally more efficient. Moreover, the alternately iterative strategy in PARAFAC provides more accurate parameter estimation. As a result, PARAFAC estimators are more appearing for multi-dimensional data analysis.

Usually, white Gaussian noise assumptions are addressed in a bistatic MIMO radar for development of estimators. In practice, however, this assumption may be occasionally violated and yield poor performances, since the noise may be spatially colored. In array signal processing, the covariance method is a powerful tool to handle spatially colored noise, especially with a sufficiently large number of snapshots. Nevertheless, the third-order PARAFAC decomposition algorithms in [8] and [9] are unable to deal with this general scenario. In this paper, the covariance tensor-based PARAFAC algorithm is devised, which can be easily expanded to the above nonideality. Unlike the covariance matrix methods, the array covariance data is rearranged into a fourth-order tensor, which coincides the PARAFAC model. A quadrilinear alternately least squares (QALS) scheme is proposed for PARAFAC decomposition. Thereafter, the LS fitting method is utilized for joint DOD and DOA estimation. The proposed method can exploit the multi-dimensional nature of the array data, and it provides closed-form solutions for DODs and DOAs, which are automatically paired. Furthermore, the uniqueness issue as well as the computational complexities are discussed. Finally, numerical simulations are given to demonstrate the effectiveness of the proposed method.

The paper outline is as follows. Necessary preliminaries of tensor decomposition and the data model for the bistatic MIMO radar is presented in section 2. The details of the proposed scheme are given in section 3. Extra extensions of the proposed algorithm are discussed in section 4. The algorithm is analyzed in section 5. Simulation results are illustrated in section 6. The paper is ended by a brief conclusion in section 7.

Notation, bold capital letters, e.g., $\mathbf{X}$, bold lowercase letters, e.g., $\mathbf{x}$, and boldface Euler script letters, e.g., $\mathcal{X}$, denote matrices, vectors, and tensors, respectively. The identity matrix is denoted by $\mathbf{I}$. The superscript $(\mathbf{X})^T$,$(\mathbf{X})^H$ and $(\mathbf{X})^{-1}$ stand for the operations of transpose, Hermitian transpose and inverse, respectively; $\otimes$ and $\odot$ represent, respectively, the Kronecker product and the Khatri-Rao product (column-wise Kronecker product); $\mathrm{diag}(\cdot)$ and $\mathrm{vec}(\cdot)$ denotes the diagonalization and the vectorization operation, respectively. $\mathrm{E}\{\cdot\}$ and $\mathrm{phase}\{\cdot\}$ return the expectation and phase of a variable, respectively. $\|\cdot\|_F$ denote the Frobenius norm of a matrix/tensor.

II. TENSOR PRELIMINARIES AND DATA MODEL

A. TENSOR PRELIMINARIES

A tensor is an N-way vector that can be regarded as the higher-order analogue of a vector. In fact, a vector is a one-way vector, a matrix is a two-way vector. Before we introduce the details of the proposed algorithm, some necessary preliminaries concerning tensor and tensor decomposition are given.

Definition 1 (Unfolding): The mode-n unfolding of an N-th order tensor $\mathcal{X} \in \mathrm{C}^{I_1\times I_2\times\cdots\times I_N}$ is denoted by $[\mathcal{X}]_{(n)}$. The $(i_1, i_2, \cdots, i_N)$-element of $\mathcal{X}$ maps to the (i_n, j)-th element of $[\mathcal{X}]_{(n)}$, where $j = 1 + \sum_{k=1,k\neq n}^{N}(i_k - 1)J_k$ with $J_k = \prod_{m=1,m\neq n}^{k-1} I_m$.

Definition 2 (Mode-n Tensor-Matrix Product): The mode-n product of an N-order tensor $\mathcal{X} \in \mathrm{C}^{I_1\times I_2\times\cdots\times I_N}$ and a matrix $\mathbf{A} \in \mathrm{C}^{J_n\times I_n}$, is denoted by $\mathcal{Y} = \mathcal{X}_{\times n}\mathbf{A}$, where $\mathcal{Y} \in \mathrm{C}^{I_1\times I_2\times\cdots\times I_{n-1}\times J_n\times I_{n+1}\times\cdots\times I_N}$ and

$$\mathcal{Y}(i_1, \cdots, i_{n-1}, j_n, i_{n+1}, \cdots, i_N) = \sum_{i_n=1}^{I_n} \mathcal{X}(i_1, i_2, \cdots, i_N)\,\mathbf{A}(j_n, i_n) \tag{1}$$

Also, the mode-n product of $\mathcal{X}$ and $\mathbf{A}$ can be expressed in tensor unfolding format as $[\mathcal{Y}]_{(n)} = \mathbf{A}\,[\mathcal{X}]_{(n)}$.

Definition 3 (Tucker Decomposition): Tucker decomposition factorizes a tensor into a set of matrices and one small core tensor. The Tucker decomposition of a Nth-order tensor $\mathcal{X} \in \mathrm{C}^{I_1\times I_2\times\cdots\times I_N}$ is given by

$$\mathcal{X} = \mathcal{G}_{\times 1}\mathbf{A}_{1\times 2}\mathbf{A}_{2\times}\cdots\mathbf{A}_N \tag{2}$$

where $\mathcal{G} \in \mathrm{C}^{R_1\times R_2\times\cdots\times R_N}$ is the core tensor, $\mathbf{A}_n \in \mathrm{C}^{I_n\times R_n}$ ($n = 1, 2, \cdots, N$, and $R_n \leq I_n$) are the factor matrices. Also, Tucker decomposition is called higher-order principal component analysis since $\mathbf{A}_n$ is the base matrix of the mode-n unfolding of $\mathcal{X}$.

Definition 4 (PARAFAC Decomposition): PARAFAC decomposition is a special case of the Tucker decomposition in Eq.(2) with $\mathcal{G}$ is a diagonal tensor, i.e., $R_1 = R_2\cdots = R_N = K$, the $(k, k, \cdots, k)$-th ($k = 1, 2, \cdots, K$) element of $\mathcal{G}$ is nonzero and zeros elsewhere. Element-wise, the PARAFAC decomposition of an N-order tensor $\mathcal{X}$ with rank-K is given by

$$\mathcal{X}(i_1, i_2, \cdots, i_N) = \sum_{i=1}^{K} g_k\mathbf{A}_1(i_1, k)\,\mathbf{A}_2(i_2, k)\cdots\mathbf{A}_N(i_N, k) \tag{3}$$

where $g_k = \mathcal{G}(k, k, \cdots, k)$. Obviously, PARAFAC decomposition factorizes a tensor into sum of rank-one vectors. In tensor unfolding format, Eq.(3) can be rewritten as

$$[\mathcal{X}]_{(n)} = \mathbf{A}_n \cdot \mathbf{G} \cdot [\mathbf{A}_{n+1} \odot \cdots \mathbf{A}_N \odot \mathbf{A}_1 \cdots \odot \mathbf{A}_{n-1}]^T \tag{4}$$

where $\mathbf{G} = \mathrm{diag}([g_1, g_2, \cdots, g_K])$.

B. DATA MODEL

Although our algorithm is suitable to an arbitrary linear array manifold, we will illustrate the idea for the uniform linear array (ULA) geometry in both the transmit array and the receive array. As shown in Fig.1, the bistatic MIMO radar system is configured with M transmit elements and N receive elements, both of which are half-wavelength spacing. Suppose there are K uncorrelated far-field targets appearing in

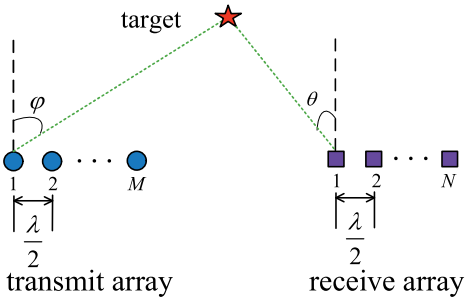

FIGURE 1. Bistatic MIMO radar configuration.

the same range, the DOD and DOA of the k-th target are φ_k and θ_k. Assume the transmit elements emit mutual orthogonal coded waveforms $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_M]^T$, where $\mathbf{w}_m \in \mathrm{C}^{Q\times 1}$ is the m-th ($l = 1, 2, \cdots, M$) baseband code with length Q and $\mathbf{w}_m^H \mathbf{w}_n = \begin{cases} Q, & m \neq n \\ 0, & m = n \end{cases}$. Additional assumption is that the Doppler frequency f_k and scattering coefficient β_k of the k-th target keep invariant during a pulse period. The noiseless received array signal at the l-th pulse duration is therefore given by

$$\mathbf{X}_l = \mathbf{A}_r \mathrm{diag}\left(\mathbf{s}_l\right) \mathbf{A}_t^T \mathbf{W} \tag{5}$$

where $\mathbf{A}_r$, $\mathbf{A}_t$ and $\mathbf{s}_l$ denote, respectively, the receive direction matrix, the transmit direction matrix and the echo coefficient vector with

$$\begin{aligned}
\mathbf{A}_r &= \left[\mathbf{a}_r\left(\theta_1\right), \mathbf{a}_r\left(\theta_2\right), \cdots, \mathbf{a}_r\left(\theta_K\right)\right] \in \mathrm{C}^{N\times K} \\
\mathbf{a}_r(\theta_k) &= \left[1, e^{j\pi \sin\theta_k}, \ldots, e^{j\pi(N-1)\sin\theta_k}\right]^T \\
\mathbf{A}_t &= \left[\mathbf{a}_t\left(\varphi_1\right), \mathbf{a}_t\left(\varphi_2\right), \cdots, \mathbf{a}_t\left(\varphi_K\right)\right] \in \mathrm{C}^{M\times K} \\
\mathbf{a}_t(\varphi_k) &= \left[1, e^{j\pi \sin\varphi_k}, \ldots, e^{j\pi(M-1)\sin\varphi_k}\right]^T \\
\mathbf{s}_l &= \left[\beta_1 e^{j2\pi l f_1/f_s}, \beta_2 e^{j2\pi l f_2/f_s}, \ldots, \beta_K e^{j2\pi l f_K/f_s}\right]^T
\end{aligned}$$

where f_s is the pulse repetition frequency. The array data is matched with $\mathbf{w}_1/Q, \mathbf{w}_2/Q \cdots, \mathbf{w}_m/Q$, respectively. Now we consider a coherent processing interval consisting L snapshots ($l = 0, 1, \cdots, L-1$). The output of the matched filters is

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \\ \vdots \\ \mathbf{Y}_N \end{bmatrix} = \begin{bmatrix} \mathbf{A}_t D_1\left(\mathbf{A}_r\right)\mathbf{S} \\ \mathbf{A}_t D_2\left(\mathbf{A}_r\right)\mathbf{S} \\ \vdots \\ \mathbf{A}_t D_N\left(\mathbf{A}_r\right)\mathbf{S} \end{bmatrix} = \left(\mathbf{A}_r \odot \mathbf{A}_t\right)\mathbf{S} \tag{6}$$

where $\mathbf{Y}_n = \mathbf{A}_t D_n\left(\mathbf{A}_r\right)\mathbf{S}$ ($n = 1, 2, \cdots, N$), $D_n\left(\mathbf{A}_r\right)$ returns a diagonal matrix with diagonal elements are the n-th row vector of $\mathbf{A}_r$. $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_L]$. The noisy covariance matrix of $\mathbf{Y}$ is then given by

$$\mathbf{R} = \mathrm{E}\left\{\mathbf{Y}\mathbf{Y}^H\right\} = \left[\mathbf{A}_r \odot \mathbf{A}_t\right]\mathbf{G}\left[\mathbf{A}_r \odot \mathbf{A}_t\right]^H + \mathbf{N} \tag{7}$$

where $\mathbf{G} = \mathrm{E}\left\{\mathbf{S}\mathbf{S}^H\right\}$, since the targets are uncorrelated, $\mathbf{G}$ is a diagonal matrix. $\mathbf{N}$ is the noisy covariance matrix. In practice, $\mathbf{R}$ can be estimated via finite snapshot through $\mathbf{R}_{est} = 1/L \sum_{l=1}^{L} \mathbf{Y}(:, l)\, \mathbf{Y}^H(:, l)$, with $\mathbf{Y}(:, l)$ denotes the l-th column of $\mathbf{Y}$. It should be noted that $\mathbf{R}$ is a Hermitian matrix. Similar to [6], $\mathbf{R}$ can be rearranged into a fourth-order tensor $\mathcal{R} \in \mathrm{C}^{M\times N\times M\times N}$ as follows

$$\mathcal{R} = \mathcal{G} \times_1 \mathbf{A}_{t\times 2} \mathbf{A}_{r\times 3} \mathbf{A}_{t\times 4}^* \mathbf{A}_r^* + \mathcal{N} \tag{8}$$

Herein, $\mathcal{R}$ is a Hermitian tensor [10] and $\mathbf{R}$ can be interpreted as the symmetric Hermitian unfolding of $\mathcal{R}$, denoting by $[\mathcal{R}]_{(H)} = \mathbf{R}$. Similarity, $[\mathcal{G}]_{(H)} = \mathbf{G}$ and $[\mathcal{N}]_{(H)} = \mathbf{N}$. Obviously, Eq.(8) presents the Tucker decomposition model for the covariance tensor.

III. THE PROPOSED ALGORITHM

A. PARAFAC MODEL

We firstly consider a ideal scenario with white Gaussian receive noise, i.e., $\mathbf{N} = \sigma^2\mathbf{I}$ is a scaled identity matrix, where σ^2 is the noise variance. As $\mathbf{G}$ is a diagonal matrix, $\mathcal{G}$ is a K-th order diagonal tensor, $\mathcal{N}$ is the symmetric Hermitian folding of $\mathbf{N}$. Therefore, the Tucker decomposition model in Eq.(8) coincides a PARAFAC decomposition formulation for the covariance tensor. To obtain the estimations of the DOD and DOA, we propose to find the best approximate of $\mathcal{R}$ as follows

$$\min \parallel \hat{\mathcal{R}} - \mathcal{R} - \mathcal{N} \parallel_F \quad st.\ \hat{\mathcal{R}} = \hat{\mathcal{G}} \times_1 \hat{\mathbf{A}}_{t\times 2} \hat{\mathbf{A}}_{r\times 3} \hat{\mathbf{A}}_{t\times 4}^* \hat{\mathbf{A}}_r^* \tag{9}$$

Since the noise is white Gaussian, the noise covariance σ^2 can be easily estimated via the eigendecomposition method, thus the noise item can be direct removed from $\mathcal{R}$ via $\left[\tilde{\mathcal{R}}\right]_{(H)} = \mathbf{R} - \hat{\sigma}^2\mathbf{I}$. Then the optimization in Eq. (9) can be reduced to solve

$$\min \parallel \hat{\mathcal{R}} - \tilde{\mathcal{R}} \parallel_F \quad st.\ \hat{\mathcal{R}} = \hat{\mathcal{G}} \times_1 \hat{\mathbf{A}}_{t\times 2} \hat{\mathbf{A}}_{r\times 3} \hat{\mathbf{A}}_{t\times 4}^* \hat{\mathbf{A}}_r^* \tag{10}$$

As a result, the estimation of DOD and DOA is linked to a quadrilinear decomposition problem of the array covariance tensor.

B. QUADRILINEAR DECOMPOSITION

According to **Definition 4**, $\hat{\mathcal{R}}$ can be expressed in matricization format as

$$\begin{cases}
\left[\hat{\mathcal{R}}\right]_{(1)} = \hat{\mathbf{A}}_t \hat{\mathbf{G}} \left[\hat{\mathbf{A}}_r \odot \hat{\mathbf{A}}_t^* \odot \hat{\mathbf{A}}_r^*\right]^T \\
\left[\hat{\mathcal{R}}\right]_{(2)} = \hat{\mathbf{A}}_r \hat{\mathbf{G}} \left[\hat{\mathbf{A}}_t^* \odot \hat{\mathbf{A}}_r^* \odot \hat{\mathbf{A}}_t\right]^T \\
\left[\hat{\mathcal{R}}\right]_{(3)} = \hat{\mathbf{A}}_t^* \hat{\mathbf{G}} \left[\hat{\mathbf{A}}_r^* \odot \hat{\mathbf{A}}_t \odot \hat{\mathbf{A}}_r\right]^T \\
\left[\hat{\mathcal{R}}\right]_{(4)} = \hat{\mathbf{A}}_r^* \hat{\mathbf{G}} \left[\hat{\mathbf{A}}_t \odot \hat{\mathbf{A}}_r \odot \hat{\mathbf{A}}_t^*\right]^T
\end{cases} \tag{11}$$

Consequently, the tensor optimization problem in Eq. (10) can be transfer into the following joint optimization problems

$$\begin{cases} \min \parallel \tilde{\mathbf{R}}_1 - \hat{\mathbf{A}}_t \hat{\mathbf{G}} \left[\hat{\mathbf{A}}_r \odot \hat{\mathbf{A}}_t^* \odot \hat{\mathbf{A}}_r^* \right]^T \parallel_F \\ \min \parallel \tilde{\mathbf{R}}_2 - \hat{\mathbf{A}}_r \hat{\mathbf{G}} \left[\hat{\mathbf{A}}_t^* \odot \hat{\mathbf{A}}_r^* \odot \hat{\mathbf{A}}_t \right]^T \parallel_F \\ \min \parallel \tilde{\mathbf{R}}_3 - \hat{\mathbf{A}}_t^* \hat{\mathbf{G}} \left[\hat{\mathbf{A}}_r^* \odot \hat{\mathbf{A}}_t \odot \hat{\mathbf{A}}_r \right]^T \parallel_F \\ \min \parallel \tilde{\mathbf{R}}_4 - \hat{\mathbf{A}}_r^* \hat{\mathbf{G}} \left[\hat{\mathbf{A}}_t \odot \hat{\mathbf{A}}_r \odot \hat{\mathbf{A}}_t^* \right]^T \parallel_F \end{cases} \tag{12}$$

where $\tilde{\mathbf{R}}_n$ $(n = 1, 2, 3, 4)$ is the mode-n unfolding of $\tilde{\mathcal{R}}$. A common technique to solve the optimization in Eq.(12) is QALS, which iteratively optimizing one of $\hat{\mathbf{G}}$, $\hat{\mathbf{A}}_t$, $\hat{\mathbf{A}}_r$, $\hat{\mathbf{A}}_t^*$ and $\hat{\mathbf{A}}_r^*$, respectively, while fixing the others. From Eq. (12) we can get the LS solutions for $\mathbf{A}_t$, $\mathbf{A}_r$, $\mathbf{A}_t^*$ and $\mathbf{A}_r^*$ are

$$\begin{cases} \hat{\mathbf{A}}_t = \tilde{\mathbf{R}}_1 \left(\hat{\mathbf{G}} \left[\hat{\mathbf{A}}_r \odot \hat{\mathbf{A}}_t^* \odot \hat{\mathbf{A}}_r^* \right]^T \right)^\dagger \\ \hat{\mathbf{A}}_r = \tilde{\mathbf{R}}_2 \left(\hat{\mathbf{G}} \left[\hat{\mathbf{A}}_t^* \odot \hat{\mathbf{A}}_r^* \odot \hat{\mathbf{A}}_t \right]^T \right)^\dagger \\ \hat{\mathbf{A}}_t^* = \tilde{\mathbf{R}}_3 \left(\hat{\mathbf{G}} \left[\hat{\mathbf{A}}_r^* \odot \hat{\mathbf{A}}_t \odot \hat{\mathbf{A}}_r \right]^T \right)^\dagger \\ \hat{\mathbf{A}}_r^* = \tilde{\mathbf{R}}_4 \left(\hat{\mathbf{G}} \left[\hat{\mathbf{A}}_t \odot \hat{\mathbf{A}}_r \odot \hat{\mathbf{A}}_t^* \right]^T \right)^\dagger \end{cases} \tag{13}$$

Subsequently, we need to estimate $\mathbf{G}$. Since $\tilde{\mathbf{R}} = [\mathbf{A}_r \odot \mathbf{A}_t]\, \mathbf{G}\, [\mathbf{A}_r \odot \mathbf{A}_t]^H$, and we have the property vec $(\mathbf{ABC}) = \left(\mathbf{C}^T \otimes \mathbf{A}\right)$ vec $(\mathbf{B})$. Accordingly, $\hat{\mathbf{G}}$ can be obtained by fitting

$$\min \parallel \text{vec}\left(\tilde{\mathbf{R}}\right) - \left(\left[\hat{\mathbf{A}}_r \odot \hat{\mathbf{A}}_t \right]^* \otimes \left[\hat{\mathbf{A}}_r \odot \hat{\mathbf{A}}_t \right] \right) \text{vec}\left(\hat{\mathbf{G}}\right) \parallel_F \tag{14}$$

Hence, we can get the LS solution for $\mathbf{G}$ via

$$\text{vec}\left(\hat{\mathbf{G}}\right) = \left(\left[\hat{\mathbf{A}}_r \odot \hat{\mathbf{A}}_t \right]^* \otimes \left[\hat{\mathbf{A}}_r \odot \hat{\mathbf{A}}_t \right] \right) \text{vec}\left(\tilde{\mathbf{R}}\right) \tag{15}$$

The iterations in Eq.(13) and Eq.(15) will repeat until converge conditions have been satisfied, e.g., the iteration number is greater than a predetermined threshold, or $\parallel \hat{\mathcal{R}}^{\text{new}} - \hat{\mathcal{R}}^{\text{old}} \parallel_F \,/\, \parallel \hat{\mathcal{R}}^{\text{old}} \parallel_F \leq 10^{-6}$, where the superscript 'new' and 'old' denote, the estimated results in current iteration and the previous iteration, respectively.

Remark 1: The proposed approach is sensitive to the initialization. QALS can be randomly initialized, which may suffer from the slow convergence. To accelerate the convergence of the proposed method, the PM algorithm [3] is adopted for initialization.

Remark 2: QALS is quite easy to implement and is guaranteed to converge, since the conditional update of any given matrix may either improve or maintain, but cannot worsen, the current fit. Global monotone convergence to (at least) a local minimum follows directly from this observation [11]. Fortunately, we experimentally found these local minimums are usually the global minimums.

C. JOINT DOD AND DOA ESTIMATION

Once QALS is accomplished, $\hat{\mathbf{A}}_r$ and $\hat{\mathbf{A}}_t$ can be obtained. Let the rank of a matrix $\mathbf{A}$ is denoted by $k_{\mathbf{A}}$. It has been proven that if

$$k_{\mathbf{A}_t} + k_{\mathbf{A}_r} + k_{\mathbf{A}_t^*} + k_{\mathbf{A}_r^*} \geq 2K + 3 \tag{16}$$

then $\mathbf{A}_t$ and $\mathbf{A}_r$ are unique up to permutation and scaling of columns [12]. The permutation and scaling effect after QALS can be formulated as:

$$\begin{cases} \hat{\mathbf{A}}_t = \mathbf{A}_t \Omega \Delta_1 + \mathbf{N}_1 \\ \hat{\mathbf{A}}_r = \mathbf{A}_r \Omega \Delta_2 + \mathbf{N}_2 \\ \hat{\mathbf{A}}_t^* = \mathbf{A}_t^* \Omega \Delta_3 + \mathbf{N}_3 \\ \hat{\mathbf{A}}_r^* = \mathbf{A}_r^* \Omega \Delta_4 + \mathbf{N}_4 \end{cases} \tag{17}$$

where Ω is a permutation matrix, Δ_n $(n = 1, 2, 3, 4)$ is a diagonal matrix containing the scale factors with $\Delta_1 \Delta_2 \Delta_3 \Delta_4 = \mathbf{I}$, $\mathbf{N}_n$ is the associated fitting noise. Define $\mathbf{P}_r = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & \pi & \cdots & (N-1)\pi \end{bmatrix}^T \in \mathrm{R}^{N \times 2}$, $\mathbf{P}_t = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & \pi & \cdots & (M-1)\pi \end{bmatrix}^T \in \mathrm{R}^{M \times K}$. Construct $\mathbf{h}_r^k = -\text{phase}\{\mathbf{a}_r(\theta_k)\}$, $\mathbf{h}_t^k = -\text{phase}\{\mathbf{a}_t(\varphi_k)\}$, and let $\mathbf{c}_r^k = [a, \sin\theta_k]^T$, $\mathbf{c}_t^k = [b, \sin\varphi_k]^T$. As $\mathbf{A}_t$ and $\mathbf{A}_r$ are Vandermonde matrices, it is easy to find that

$$\begin{cases} \mathbf{P}_r \mathbf{c}_r^k = \mathbf{h}_r^k \\ \mathbf{P}_t \mathbf{c}_t^k = \mathbf{h}_t^k \end{cases} \tag{18}$$

Worthnoting is that the phase of $\hat{\mathbf{A}}_r$ and $\hat{\mathbf{A}}_t$ are still linear. Let $\hat{\mathbf{a}}_r(\theta_k)$ and $\hat{\mathbf{a}}_r(\theta_k)$ are the k-th columns of $\hat{\mathbf{A}}_r$ and $\hat{\mathbf{A}}_t$, respectively. Let $\hat{\mathbf{h}}_r^k = -\text{phase}\left\{\hat{\mathbf{a}}_r(\theta_k)\right\}$, $\hat{\mathbf{h}}_t^k = -\text{phase}\left\{\hat{\mathbf{a}}_t(\varphi_k)\right\}$. Then, we can get

$$\begin{cases} \hat{\mathbf{c}}_r^k = \mathbf{P}_r^\dagger \hat{\mathbf{h}}_r \\ \hat{\mathbf{c}}_t^k = \mathbf{P}_t^\dagger \hat{\mathbf{h}}_t \end{cases} \tag{19}$$

Clearly, the second elements $\hat{\mathbf{c}}_r^k(2)$ and $\hat{\mathbf{c}}_t^k(2)$ of $\hat{\mathbf{c}}_r^k$ and $\hat{\mathbf{c}}_t^k$, respectively, stand for the LS solutions for $\sin\theta_k$ and $\sin\varphi_k$. As a result, joint DOD and DOA estimation can be obtained via

$$\begin{cases} \hat{\theta}_k = \arcsin \hat{\mathbf{c}}_r^k(2) \\ \hat{\varphi}_k = \arcsin \hat{\mathbf{c}}_t^k(2) \end{cases} \tag{20}$$

Remark 3: According to Eq.(17), $\hat{\mathbf{A}}_r$ and $\hat{\mathbf{A}}_t$ share the same permutation, $\hat{\theta}_k$ and $\hat{\varphi}_k$ are therefore paired automatically.

Finally, we have achieve the proposal of the proposed QALS method, which is summarized in Table 1.

IV. EXTENSION TO SCENARIO WITH SPATIALLY COLORED NOISE

In the presence of unknown spatially colored noise, $\mathbf{N}$ is a block diagonal matrix instead of a scaled identity matrix. As a result, the denoising method though eigendecomposition will fail to work and yields degraded estimation performance. Several powerful covariance approaches have been

TABLE 1. Algorithmic steps of the QALS method.

step	operation
1	Compute $\mathbf{R}_{est} = 1/L\sum_{l=1}^{L}\mathbf{Y}(:,l)\mathbf{Y}^H(:,l)$. Estimate σ^2 through eigendecomposition, and calculate $\tilde{\mathbf{R}} = \mathbf{R}_{est} - \sigma^2\mathbf{I}$.
2	set $\hat{\mathbf{G}}$ to a identity matrix, initialize $\hat{\mathbf{A}}_t$ and $\hat{\mathbf{A}}_r$ with PM algorithm.
3	Repeat the iterations in Eq.(13) and Eq.(15) until convergence.
4	Output the DODs and DOAs through Eq.(19)-Eq.(20).

developed to eliminate the spatially colored noise. Typical denoising frameworks including spatial cross-correlation methods [13]–[15] and temporal cross-correlation methods [16], [17]. It has been shown that the temporal cross-correlation algorithms provide more accurate direction estimation than the spatial cross-correlation methods [17], since the temporal cross-correlation algorithms would not bring any virtual aperture loss. In temporal cross-correlation methods, two data matrices $\mathbf{Z}_1$ and $\mathbf{Z}_2$ formed by choosing the first $L-1$ columns and the last $L-1$ columns of $\mathbf{Y}$, i.e.,

$$\begin{cases}\mathbf{Z}_1 = (\mathbf{A}_r \odot \mathbf{A}_t)\mathbf{S}_1 + \mathbf{E}_3 \\ \mathbf{Z}_2 = (\mathbf{A}_r \odot \mathbf{A}_t)\mathbf{S}_2 + \mathbf{E}_4\end{cases} \tag{21}$$

where $\mathbf{S}_1$ and $\mathbf{S}_2$ denote the first and last $L-1$ columns of $\mathbf{S}$, $\mathbf{E}_3$ and $\mathbf{E}_4$ denote the first and last $L-1$ columns of the matched noise $\mathbf{E}$, respectively. Thereafter, the cross-correlation matrix $\mathbf{R}_z$ is constructed as

$$\mathbf{R}_z = \mathrm{E}\left\{\mathbf{Z}_1\mathbf{Z}_2^H\right\} = [\mathbf{A}_r \odot \mathbf{A}_t]\,\mathbf{G}_1\,[\mathbf{A}_r \odot \mathbf{A}_t]^H \tag{22}$$

where $\mathbf{G}_1 = \mathrm{E}\left\{\mathbf{S}_1\mathbf{S}_2^H\right\}$, which is a diagonal matrix in the presence of uncorrelated targets. The colored noise is suppressed due to $\mathrm{E}\left[\mathbf{S}_1\mathbf{S}_2^H\right] = \mathbf{0}$. Thereafter, the subspace approaches can be direct applied. Similarity, $\mathbf{R}_z$ can be rearranged into a fourth-order tensor as

$$\mathcal{R}_z = \mathcal{G}_{1\times 1}\mathbf{A}_{t\times 2}\mathbf{A}_{r\times 3}\mathbf{A}^*_{t\times 4}\mathbf{A}^*_r \tag{23}$$

where $[\mathcal{R}_z]_{(H)} = \mathbf{R}_z$ and $[\mathcal{G}_1]_{(H)} = \mathbf{G}_1$. Note the model in Eq.(23) concides the PARAFAC decomposition model, the proposed QALS method (using steps 2-4) is applicable.

Remark 4: Actually, white Gaussian noise is a special case of the colored noise, thus the extened algorithm also suitable for white Gaussian noise scenario.

V. DISCUSSION

A. ALGORITHM IDENTIFIABILITY

Generally, the low-rank approximation is non-unique while low-rank tensor factorization is essentially unique under mild conditions. Eq.(16) implies the sufficient condition for the uniqueness of PARAFAC decomposition. Usually, we have $L \gg M$ and $L \gg N$, thus $k_{\mathbf{A}_t} = k_{\mathbf{A}_t^*} = M$, $k_{\mathbf{A}_r} = k_{\mathbf{A}_r^*} = N$. Therefore, the proposed algorithm can idendity up to $M+N-1$ targets.

TABLE 2. Comparison of the complexity.

Method	Main computation load
ESPRIT	$\mathcal{O}\left(M^3N^3\right)$
HOSVD	$\mathcal{O}\left(M^3N^3\right)$
PARAFAC	$\mathcal{O}\left(MNLK + MLK^2 + NLK^2 + MNK^2\right)$
QALS	$\mathcal{O}\left(M^2NK^2 + MN^2K^2 + M^2N^2K^4\right)$

B. COMPUTATIONAL COMPLEXITY

The main computational load of the proposed algorithm is the QALS in Eq.(13), which requires $l\left[3M^2NK^2 + 3MN^2K^2 + 3M^2N^2K^4\right]$ complex multiplications, where l denotes the number of iterations. Experimentally, QALS converges after no more than 20 iterations. Table 2 lists the main complexity of the some typical closed-form-based algorithms. The complexity of ESPRIT in [2] is mainly concentrated in the eigendecomposition of the array covariance matrix, which is on the order $\mathcal{O}\left(M^3N^3\right)$. The complexity of the higher-order singular value decomposition (HOSVD) in [6] is four times of that in ESPRIT, but they are on the same order. The iteration of the PARAFAC in [8] has a complexity of $\mathcal{O}\left(MNLK + MLK^2 + NLK^2 + MNK^2\right)$. Comparing with the ESPRIT and HOSVD approaches, the proposed method is computationally more efficient in the presence of Massive MIMO systems. However, the proposed method may has higher complexity than PARAFAC, but it should be noticed that the proposed method is more flexible than PARAFAC, which will be shown in the simulation section.

VI. SIMULATION RESULTS

In this section, 200 Monte Carlo trials are applied to verify the significance of the proposed method. Simulation conditions are set to $K = 3$ targets with DODs and DOAs are $(\theta_1, \varphi_1) = (15°, 10°)$, $(\theta_2, \varphi_2) = (-25°, -20°)$ and $(\theta_3, \varphi_3) = (35°, 30°)$, respectively. The associate Doppler frequencies are 100, 500 and 800Hz, and the pulse repetition frequency is set to $f_s = 20$ KHz. The scattering coefficients satisfy Swerling I model. The bistatic MIMO radar is equipped with M transmit elements and N receive elements. Pulse number Q and pulse repeat frequency f_s are set to $Q = 256$ and L snapshots are collected. The transmit baseband coded matrix is consist of the first M rows of a $Q \times Q$ Hadamard matrix. In the simulations, the signal-to-noise ratio is defined as the ratio of signal power to noise power noise before matched filtering. Two measures are used for performance assessment [17]. One is the root mean square error (RMSE) defined as

$$\mathrm{RMSE} = \frac{1}{K}\sum_{k=1}^{K}\sqrt{\frac{1}{200}\sum_{i=1}^{200}\left\{\left(\hat{\theta}_{i,k} - \theta_k\right)^2 + \left(\hat{\varphi}_{i,k} - \varphi_k\right)^2\right\}}$$

where $\hat{\theta}_{i,k}$ and $\hat{\varphi}_{i,k}$ represent the estimates of θ_k and φ_k for the i-th trial. The other one is the probability of the successful detection(PSD), in which a successful trial is recognized if the absolute error of all the estimated angles are smaller than 0.2°.

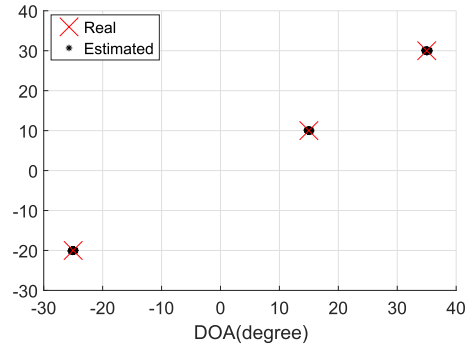

FIGURE 2. Scatter results of the proposed algorithm.

In the first simulation, we test the scatter results of the proposed in the presence of white Gaussian noise, other conditions are set to $M = 8$, $N = 6$ and $L = 200$. Fig.2 shows the results at SNR= 15dB, from which we can observe that the DODs and DOAs of the three targets can be precisely estimated and correctly paired.

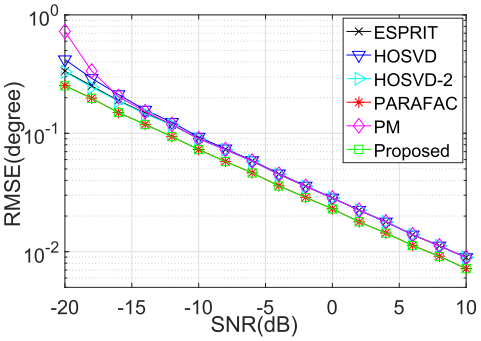

FIGURE 3. RMSE comparison versus SNR with white Gaussian noise.

In the second simulation, we examine the RMSE and PSD performances of the proposed method versus SNR in the presence of white Gaussian noise, where $M = 8$, $N = 6$ and $L = 200$ are considered. For comparison, results of the ESPRIT algorithm [2], the PM method [3], the PARAFAC scheme [8], the direct HOSVD approach(marked with HOSVD) as well as the covariance tensor-based HOSVD (marked with HOSVD2) [6] are included. Fig.3 illustrates the RMSEs comparison. As expected, the tensor-based estimators have more accurate directions estimation performances than the matrix-based estimators at low SNR regions, which benefit from the tensor gain. Another interesting observation is that the PARAFAC estimators outperform the other estimators, since the LS strategy in PARAFAC takes full advantage of the array degree-of-freedoms (DOF) while DOF of the rotation invariant technique in the other estimators is normally lossy. Fig.4 gives the PSD curves. As shown in Fig.4, all the methods exhibit a 100% successful detection at high SNR regions. When the SNR decreases, the PSD of each method starts to drop at a certain point, which is known as the SNR threshold. Clearly, the PARAFAC estimators provide lower SNR thresholds than the others. It is easy to find from Fig.3 and Fig.4 that the proposed QALS method

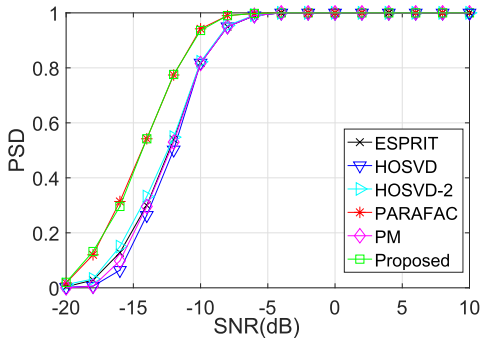

FIGURE 4. PSD comparison versus SNR with white Gaussian noise.

has comparable estimation performance to the PARAFAC algorithm when faced with white Gaussian noise.

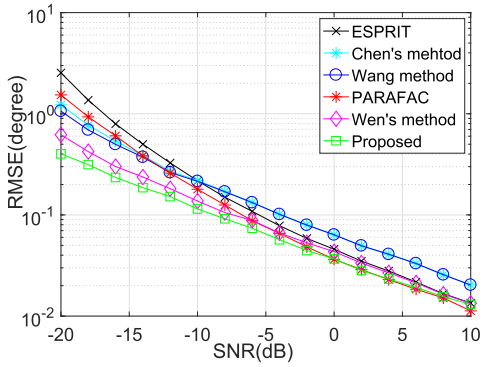

FIGURE 5. RMSE comparison versus SNR with spatially colored noise.

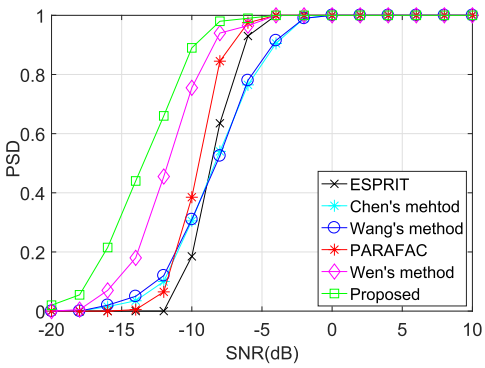

FIGURE 6. PSD comparison versus SNR with spatially colored noise.

In the third simulation, the noise is assumed to be spatially colored which is modeled as a second-order autoregressive (AR) process with the coefficients $z = [1, -1, 0.9]$. The the RMSE and PSD performances of the proposed method is examined by varying the SNR, where $M = 8$, $N = 6$ and $L = 200$ are considered. Moreover, we compare the proposed with ESPRIT and PARAFAC. Besides, the denoising methods in [13] (marked with Chen's method), [15] (Marked with Wang's method) and [17] (Marked with Wen's method) are included. Fig.5 and Fig.6 present the RMSEs and the PSDs, respectively. It can be observed that the ESPRIT method and the PARAFAC method provide the worst performance at low SNR regions, since they may fail to work in the presence of colored noise. Due to the virtual aperture loss, Chen's method and Wang's method will have worse performance

than ESPRIT and PARAFAC. Since the proposed method and Wen's method do not hurt the array aperture, they have much better performance than the others at low SNR regions. Furthermore, the proposed method is more appealing from the perspective of estimation accuracy, as the iteration framework in QALS obtains a more accurate parameter estimation than the Tucker decomposition in Wen's method.

VII. CONCLUSION

In this paper, a covariance tensor-based PARAFAC algorithm has been proposed for joint DOD and DOA estimation in bistatic MIMO radar. The covariance matrix of the matched array data is arranged into a fourth-order tensor, which coincides the PARAFAC decomposition model. A quadrilinear decomposition framework is developed to iterative fit the PARAFAC model. Finally, joint DODs and DOAs are obtained by utilizing the LS criterion. The proposed method can be viewed as the covariance version of the PARAFAC algorithm, and it can be easily extended to the scenario with spatially colored noise. Simulation results indicate the proposed method provides comparable or better estimation performance than the existing methods, which will lead to a brighter prospect in applications.

REFERENCES

[1] X. F. Zhang, L. Y. Xu, L. Xu, and D. Z. Xu, "Direction of departure (DOD) and direction of arrival (DOA) estimation in MIMO radar with reduced-dimension MUSIC," *IEEE Commun. Lett.*, vol. 14, no. 12, pp. 1161–1163, Dec. 2010.

[2] C. Duofang, C. Baixiao, and Q. Guodong, "Angle estimation using ESPRIT in MIMO radar," *Electron. Lett.*, vol. 44, no. 12, pp. 770–771, Jun. 2008.

[3] J. Li, X. Zhang, and H. Chen, "Improved two-dimensional DOA estimation algorithm for two-parallel uniform linear arrays using propagator method," *Signal Process.*, vol. 92, no. 12, pp. 3032–3038, Dec. 2012.

[4] B. Tang, J. Tang, Y. Zhang, and Z. Zheng, "Maximum likelihood estimation of DOD and DOA for bistatic MIMO radar," *Signal Process.*, vol. 93, no. 5, pp. 1349–1357, 2013.

[5] J. P. C. L. da Costa, F. Roemer, M. Haardt, and R. T. de Sousa, Jr., "Multi-dimensional model order selection," *EURASIP J. Adv. Signal Process.*, vol. 26, Dec. 2011, Art. no. 2011.

[6] Y. Cheng, R. Yu, H. Gu, and W. Su, "Multi-SVD based subspace estimation to improve angle estimation accuracy in bistatic MIMO radar," *Signal Process.*, vol. 93, pp. 2003–2009, Jul. 2013.

[7] X. Wang, W. Wang, J. Liu, Q. Liu, and B. Wang, "Tensor-based real-valued subspace approach for angle estimation in bistatic MIMO radar with unknown mutual coupling," *Signal Process.*, vol. 116, pp. 152–158, Nov. 2015.

[8] X. Zhang, Z. Xu, L. Xu, and D. Xu, "Trilinear decomposition-based transmit angle and receive angle estimation for multiple-input multiple-output radar," *IET Radar Sonar Navigat.*, vol. 5, no. 6, pp. 626–631, Jul. 2011.

[9] B. Xu, Y. Zhao, Z. Cheng, and H. Li, "A novel unitary PARAFAC method for DOD and DOA estimation in bistatic MIMO radar," *Signal Process.*, vol. 138, pp. 273–279, 2017.

[10] M. Haardt, F. Roemer, and G. D. Galdo, "Higher-order SVD-based subspace estimation to improve the parameter estimation accuracy in multi-dimensional harmonic retrieval problems," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3198–3213, Jul. 2008.

[11] 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.

[12] 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.

[13] J. Chen, H. Gu, and W. Su, "A new method for joint DOD and DOA estimation in bistatic MIMO radar," *Signal Process.*, vol. 90, no. 2, pp. 714–718, 2010.

[14] H. Jiang, J.-K. Zhang, and K. M. Wong, "Joint DOD and DOA estimation for bistatic MIMO radar in unknown correlated noise," *IEEE Trans. Veh. Technol.*, vol. 64, no. 11, pp. 5113–5125, Nov. 2015.

[15] X. Wang, W. Wang, X. Li, and J. Wang, "A tensor-based subspace approach for bistatic MIMO radar in spatial colored noise," *Sensors*, vol. 14, no. 3, pp. 3897–3907, 2014.

[16] W.-B. Fu, T. Z. Su, Y.-B. Zhao, and X.-H. He, "Joint estimation of angle and doppler frequency for bistatic mimo radar in spatial colored noise based on temporal-spatial structure," *J. Electron. Inf. Technol.*, vol. 33, no. 7, pp. 1649–1654, 2011.

[17] F. Wen, X. Xiong, J. Su, and Z. Zhang, "Angle estimation for bistatic MIMO radar in the presence of spatial colored noise," *Signal Process.*, vol. 134, pp. 261–267, May 2017.

ZIYI WANG was born in Huanggang, China, in 1999. She is currently a Junior majoring in automation with the Electronic and Information School, Yangtze University, China. She joined the Signal Processing Lab, Yangtze University, in 2016. Her research interests include artificial intelligence, wireless communications, and array signal processing. She received the Innovation and Entrepreneurship Fund for College Students from Yangtze University in 2016.

CHANGXIN CAI was born in Jingzhou, China, in 1974. He received the B.S. degree in the electronic instrument and measurement technology and the M.S. degree in geodetection and information technology from Jianghan Petroleum University, Jingzhou, China, in 1998 and 2003, respectively, and the Ph.D. degree in petroleum and natural gas engineering from Yangtze University, China, in 2013. Since 2003, he has been with the Electronic and Information School, Yangtze University, where he is currently an Associate Professor, a Master Instructor, and an Associate Dean, mainly focusing on signal and information processing, intelligent instrument design, and measurement and control of oil and gas fields.

FANGQING WEN (M'17) was born in Yangzhou, China, in 1988. He received the B.S. degree in electronic engineering from the Hubei University of Automotive Technology, Shiyan, China, 2011, the master's degree from the College of Electronics and Information Engineering, Nanjing University of Aeronautics and Astronautics (NUAA), China, in 2013, and the Ph.D. degree from NUAA, in 2016. From 2015 to 2016, he was a Visiting Scholar with the University of Delaware, USA. Since 2016, he has been with the Electronic and Information School, Yangtze University, China, where he is currently an Assistant Professor. His research interests include multiple-input multiple-output radar, array signal processing, and compressive sensing. He is member of the Chinese Institute of Electronics.

DONGMEI HUANG was born in Yangzhou, China, in China, in 1986. She received the B.S. degree from the Naval Command College, Nanjing, in 2010, and the M.S. degree from the Nanjing University of Aeronautics and Astronautics, China, in 2015. She is currently a Lecturer with the Department of Training Management, Naval Command College. Her research interests include array signal processing, radar signal processing, and tensor signal processing.