IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 18, 2019, accepted July 19, 2019, date of publication July 23, 2019, date of current version August 8, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2930531

Off-Grid DOA Estimation for Colocated MIMO Radar via Reduced-Complexity Sparse Bayesian Learning

TINGTING LIU[1,2], FANGQING WEN[3], (Member, IEEE), LEI ZHANG[3], AND KE WANG[3]

[1]School of Economics and Management, Yangtze University, Jingzhou 434023, China
[2]State Key Laboratory of Marine Resource Utilization in South China Sea, Hainan University, Haikou 570228, China
[3]National Demonstration Center for Experimental Electrical and Electronic Education, Yangtze University, Jingzhou 434023 China

Corresponding author: Tingting Liu (516112@yangtzeu.edu.cn)

This work was supported by the National Natural Science Foundation of China under Grant 61701046 and Grant 61871218.

ABSTRACT Recent advance on signal processing has witnessed increasing interest in machine learning. In this paper, we revisit the problem of direction-of-arrival (DOA) estimation for colocated multiple-input multiple-output (MIMO) radar from the perspective of machine learning. The reduced-complexity transformation is first applied on the array data from matched filters, thus eliminating the redundancy of the array data for the relief of calculational burden. Furthermore, the pre-whitening is followed to obtain a simplified noise model. Finally, the DOA estimation is linked to off-grid sparse Bayesian learning (OGSBL), which does not require to update the noise hyper-parameter, and a block hyper-parameter is utilized to accelerate the convergence of the OGSBL algorithm. The proposed estimator provides better DOA estimation accuracy than the existing peak searching algorithm. The effectiveness of the proposed algorithm is verified via numerical simulation.

INDEX TERMS Array signal processing, MIMO radar, DOA estimation, off-grid, sparse Bayesian learning.

I. INTRODUCTION

The past decade has witnessed an explosive growth in machine learning. As one of the most important subset of artificial intelligence, machine learning enables the computer system to perform a specific task efficiently, while only patterns and inference are utilized. Usually, machine learning is interpreted as a comprehensive subject that mix data mining, optimization and statistics. Owing to its potential prospect, machine learning has been extensively used in medical science [1], internet-of-thing [2], wireless communications [3]–[7] and image processing [8], [9].

Direction-of-arrival (DOA) estimation is one of the basic tasks in colocated multiple-input multiple-output (MIMO) radar. Up to now, thousands of algorithms have been reported in this topic. Typical estimation algorithms including multiple signal classification (MUSIC) [10], [11], estimation of signal parameters via rotational invariance techniques (ESPRIT) [12], [13], and tensor-based methods [14]–[19]. Besides, many efforts have been devoted to direction finding in the presence of sensor errors [20]–[25].

The associate editor coordinating the review of this manuscript and approving it for publication was Derek Abbott.

Generally, ESPRIT can obtain closed-form solutions for DOA estimation at the cost of a few computationally load, while the tensor approaches provides more accurate DOA estimation performance with higher complexity, as the tensor nature can be explored. The spectrum grid search methods (such as MUSIC) have drawn extensive attention in the past decades since they often provide more accurate DOA estimation performance than ESPRIT, as more degree-of-freedom (DOF) can be exploited by the former. In the peak search methods, a grid must be initialized by the algorithm. However, it is often very hard to choose a 'proper' grid. As it is well known, a refined grid leads to accurate estimation performance but the estimator may suffers from exhaustive peak search. On the contrary, a sparse grid is helpful to lower the computational load, but it is quite possible to obtain imprecise DOA estimation, as the true DOAs are off-the-grid with high probability. To obtain off-grid DOA estimation, the perturbed sparse model is established by exploiting coarse grid as dictionary matrix and some strategies have been put forward to recovery the sparse support [26]–[30]. In [26], the perturbations (the intervals between true DOAs and the nearest grids) is approximated by the first-order Taylor series, and a sparse total least squares algorithm is presented.

An off-grid sparse Bayesian learning (OGSBL) framework is investigated in [27], where the off-grid gaps are considered to fulfill the uniform distribution. An iterative framework is proposed in [28], which is suitable to solve the off-grid DOA estimation problem form the one-bit measurement. In [29], an improved Bayesian inference method is derived, in which the root method replace the evidence procedure in [27] to update the grid more efficiently. Besides, a covariance matrix-based Bayesian learning algorithm is introduced in [30], the block sparse model is formulated, which taken the temporal correlation into consideration. In [31], an off-grid solver is developed for monostatic MIMO radar. Unlike the methods in [26]–[30], it is mainly focus on the off-grid DOA estimation problem in the presence of unknown mutual coupling.

Although the above mentioned algorithms in [27], [29], [30] are suitable to solve the off-grid problem, they are unsuitable for colocated MIMO radar directly as they involve huge calculation. In the presence of uniform linear array (ULA) geometries, the array measurements from matched filters of colocated MIMO radar are redundant. Fortunately, there are some reduce-dimension strategies to reduce the complexity of the multi-dimensional measurement. For instance, the reduced-dimension (RD) transform scheme [10], [11], [32], the reduced-complexity (RC) framework [33], the unitary transform methodology [18]. Usually, the RD method is utilized to transform a multi-parameters optimization problem into a quadratic problem, while the unitary transform is adopted to change a complex-value problem into a real-value problem. In contrast to the RD method and the unitary transform approach, RC method try to remove the redundancy of the measurement, which coincides the ULA-based colocated MIMO radar. Several RC algorithms have been developed for DOA estimation in monostatic MIMO radar, e.g., RC-ESPRIT [33], RC-MUSIC [34] and RC-based optimization method [35]. In [36], a RC-OGSBL algorithm is presented, which extend the OGSBL framework to colocated MIMO configuration. Unfortunately, the RC operation in which will cause the nonuniform noise problem, thus degrades the estimation performance.

In this paper, we revisit DOA estimation problem in colocated MIMO radar, and a novel RC-OGSBL estimator is proposed. The main contributions of this paper are listed as follows:

- The RC operation is applied in colocated MIMO radar to decrease the computational load. Unlike the RC transformation in [36], nonuniform noise would not appear in the proposed estimator.
- A pre-whitening model is constructed via the covariance measurement. By exploiting the covariance matrix model, the dimension of the model is reduced, and the noise variance is normalized to one.
- An off-grid framework is derived. Similar to [30], the temporal correlation of sparse measurements are explored. Moreover, the dictionary in sparse recovery can be adapted with the algorithm.
- Numerical simulations are designed to exam the estimation performance of the proposed algorithm.

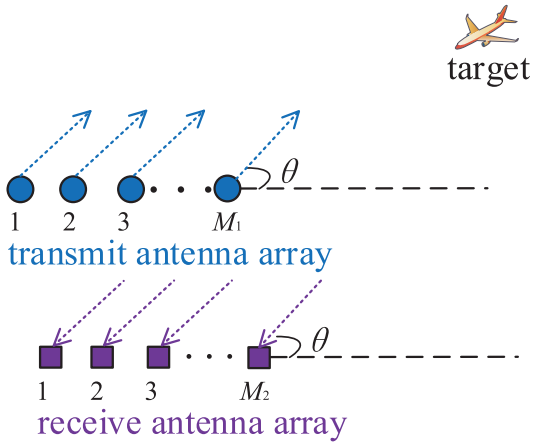


FIGURE 1. Illustration of monostatic MIMO radar.

Compared with the OGSBL algorithm in [27], the proposed algorithm has the following advantages: a). It has much lower complexity as the redundancy of the measurement has been removed via the RC transformation. b). It does not need to estimate the noise variance, thus it decreases the complexity in the iteration. Although the proposed algorithm provides very close estimation accuracy to the OGSBL algorithm, it is computationally more efficient, thus it is superior than the OGSBL algorithm.

This paper is organized as follows. The data model for DOA estimation in colocated MIMO radar is given in section II. The proposed estimator is elaborated in section III. Numerical simulations and discussions are provided in section IV. Finally, the paper is ended with a brief conclusion in section V.

II. SIGNAL MODEL

Let us consider a narrowband monostatic MIMO radar model, as illustrated in Fig. 1. The radar system is equipped with M_1 transmit elements and M_2 receive elements, both of which are one dimensional ULAs. Suppose that the transmit array emit M_1 normalized orthogonal pulse waveforms $\{p_{m_1}(t)\}_{m_1=1}^{M_1}$, i.e.,

$$\int_{T_p} p_{m_1}(t)\, p_{m_2}^*(t)\, dt = \delta(m_1 - m_2), \tag{1}$$

where t denotes the fast time index (time index during a radar pulse), T_p accounts for the pulse duration, $(\cdot)^*$ denotes the conjugate operation, $\delta(\cdot)$ is the Kronecker delta. Assume that there are K far field targets, the reflected echoes received by the receive antenna array can be shown as

$$\mathbf{r}(t,\tau) = \sum_{k=1}^{K} s_k(\tau)\, \mathbf{a}_r(\theta_k)\, \mathbf{a}_t^T(\theta_k)\, \mathbf{p}(t) + \mathbf{w}(t,\tau), \tag{2}$$

where τ is the slow time index (pulse index), $s_k(\tau)$ stands for the associate reflection coefficient, θ_k is the DOA of the k-th ($k = 1, 2, \cdots, K$) target, $\mathbf{a}_t(\theta_k) \in \mathbb{C}^{M_1 \times 1}$ and $\mathbf{a}_r(\theta_k) \in \mathbb{C}^{M_2 \times 1}$ are the transmit steering vector and the receive steering vector corresponding to the k-th target, respectively, $\mathbf{p}(t) = \left[p_1(t), p_2(t), \cdots, p_{M_1}(t)\right]^T$ is the waveform vector. $\mathbf{w}(t, \tau)$ is the zero mean Gaussian noise vector with variance is σ^2 i.e.,

$$\mathrm{E}\left\{\mathbf{w}(t_1, \tau)\, \mathbf{w}^H(t_2, \tau)\right\} = \sigma^2 \mathbf{I} \cdot \delta(t_1 - t_2), \tag{3}$$

where $\mathrm{E}\{\cdot\}$ returns the expectation of a variable. The contribution of the m_1-th ($m_1 = 1, 2, \cdots, M_1$) entity to $\mathbf{a}_t(\theta_k)$ and the contribution of m_2-th ($m_2 = 1, 2, \cdots, M_2$) entity of $\mathbf{a}_r(\theta_k)$ are

$$a_{m_1}(\theta_k) = \exp\{-j2\pi(m_1 - 1) d\cos(\theta_k)/\lambda\}, \tag{4}$$
$$a_{m_2}(\theta_k) = \exp\{-j2\pi(m_2 - 1) d\cos(\theta_k)/\lambda\}, \tag{5}$$

where d is the inter-element distance, λ is the carrier wavelength. Next, $\mathbf{r}(t, \tau)$ is matched with $\mathbf{p}(t)$ and yields

$$\begin{aligned}\mathbf{x}(\tau) &= \mathrm{vec}\left(\int_{T_p} \mathbf{r}(t, \tau)\, \mathbf{p}^H(t)\, dt\right) \\ &= \sum_{k=1}^{K} \left[\mathbf{a}_t(\theta_k) \otimes \mathbf{a}_r(\theta_k)\right] s_k(\tau) + \mathbf{e}(\tau) \\ &= \mathbf{A}\mathbf{s}(\tau) + \mathbf{e}(\tau),\end{aligned} \tag{6}$$

where $\mathrm{vec}(\cdot)$ denotes the vectorization operation, the superscript $(\cdot)^H$ denotes the Hermitian transpose, $\mathbf{e}(\tau) = \mathrm{vec}\left(\int_{T_p} \mathbf{w}(t, \tau)\, \mathbf{p}^H(t)\, dt\right)$ is the matched noise vector. $\mathbf{A} = [\mathbf{a}_t(\theta_1) \otimes \mathbf{a}_r(\theta_1), \mathbf{a}_t(\theta_2) \otimes \mathbf{a}_r(\theta_2), \cdots, \mathbf{a}_t(\theta_K) \otimes \mathbf{a}_r(\theta_K)] \in \mathbb{C}^{M_1 M_2 \times K}$ is the visual response matrix, $\mathbf{s}(\tau) = [s_1(\tau), s_2(\tau), \cdots, s_K(\tau)]^T$ is the reflection coefficient vector. It has been proven in [32] that $\mathbf{n}(\tau)$ is still a Gaussian vector with zero mean and variance σ^2, i.e.,

$$\mathbf{R}_n = \mathrm{E}\left\{\mathbf{n}(\tau)\, \mathbf{n}^H(\tau)\right\} = \sigma^2 \mathbf{I} \tag{7}$$

III. THE PROPOSED ALGORITHM

The DOF (size of the effective sensor array) is very important in MIMO radar. Since there are only $M_1 + M_2 - 1$ distinct elements in each column of $\mathbf{A}$, thus $\mathbf{A}$ is redundant. In this paper, we try to estimate the DOA form $\mathbf{x}(\tau)$ via OGSBL framework. However, signal processing based on the redundant signal $\mathbf{x}(\tau)$ involves extensive computation, some measures can be taken to reduce the burden. To this end, the reduced-dimension transform and the pre-whiting processing are applied successively.

A. REDUCED-DIMENSION TRANSFORM

Firstly, the following reduced-dimension transform matrix $\mathbf{C} \in \mathbb{C}^{M_1 M_2 \times (M_1 + M_2 - 1)}$ is defined

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \end{bmatrix} \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} M_1$$
$$\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & \cdots & 0 \end{bmatrix} \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} M_1$$
$$\begin{bmatrix} \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} M_1 \tag{8}$$

One can easily find that

$$\mathbf{A} = \mathbf{C}\mathbf{B}, \tag{9}$$

where $\mathbf{B} = [\mathbf{b}(\theta_1), \mathbf{b}(\theta_2), \cdots, \mathbf{b}(\theta_K)] \in \mathbb{C}^{N \times K}, N \triangleq M_1 + M_2 - 1$. The contribution of the n-th ($n = 1, 2, \cdots, N$) entity to $\mathbf{b}(\theta_k)$ is

$$b_n(\theta_k) = \exp\{-j2\pi(n - 1) d\cos(\theta_k)/\lambda\}, \tag{10}$$

Although $\mathbf{C}$ is effective to eliminate the redundancy of $\mathbf{A}$, left multiply $\mathbf{x}(\tau)$ with $\mathbf{C}^H$ will result in nonuniform noise. To avoid noise imperfectly, a weight matrix $\mathbf{F}$ is defined [33], [34]

$$\begin{aligned}\mathbf{F} &= \mathbf{C}^H\mathbf{C} \\ &= \mathrm{diag}\left(1, 2, \cdots, \underbrace{M, \cdots, M}_{|M_1 - M_2| + 1}, \cdots, 2, 1\right) \\ &= \mathbf{F}^H,\end{aligned} \tag{11}$$

where $\mathrm{diag}(\cdot)$ returns a diagonal matrix with the diagonal elements are the entities in the brackets, $M \triangleq \min(M_1, M_2)$. Thereafter, multiplying $\mathbf{x}(\tau)$ with $\mathbf{F}^{-1/2}\mathbf{C}^H$ yields

$$\begin{aligned}\mathbf{y}(\tau) &= \mathbf{F}^{-1/2}\mathbf{C}^H\mathbf{x}(\tau) \\ &= \mathbf{F}^{-1/2}\mathbf{C}^H\mathbf{A}\mathbf{s}(\tau) + \mathbf{F}^{-1/2}\mathbf{C}^H\mathbf{e}(\tau) \\ &= \mathbf{F}^{-1/2}\mathbf{C}^H\mathbf{C}\mathbf{B}\mathbf{s}(\tau) + \mathbf{F}^{-1/2}\mathbf{C}^H\mathbf{e}(\tau) \\ &= \underbrace{\mathbf{F}^{1/2}\mathbf{B}\mathbf{s}(\tau)}_{\text{signal}} + \underbrace{\mathbf{F}^{-1/2}\mathbf{C}^H\mathbf{e}(\tau)}_{\text{noise}}\end{aligned} \tag{12}$$

Now we focus on the noise counterpart, the covariance of which is

$$\mathbf{F}^{-1/2}\mathbf{C}^H\mathrm{E}\left\{\mathbf{e}(\tau)\, \mathbf{e}^H(\tau)\right\}\mathbf{C}\mathbf{F}^{-1/2} = \sigma^2\mathbf{I} \tag{13}$$

Eq. (13) reveals that the noise after Reduced-dimension transform is still Gaussian with invariant statistical characteristics. In the presence of uncorrelated targets, i.e.,

$$\mathbf{R}_s = \mathrm{E}\left\{\mathbf{s}(\tau)\,\mathbf{s}^H(\tau)\right\} = \mathrm{diag}(\beta_1, \beta_2, \cdots, \beta_K), \quad (14)$$

where β_k $(k = 1, 2, \cdots, K)$ is the variance of the reflect coefficients associate to the k-th target. Consequently, the covariance matrix of $\mathbf{y}(\tau)$ can be expressed as

$$\begin{aligned}\mathbf{R}_y &= \mathrm{E}\left\{\mathbf{y}(\tau)\,\mathbf{y}^H(\tau)\right\} \\ &= \mathbf{F}^{1/2}\mathbf{B}\mathbf{R}_s\mathbf{B}^H\mathbf{F}^{1/2} + \sigma^2\mathbf{I}\end{aligned} \quad (15)$$

In practice, L samples $\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_L$ are available, where $\mathbf{y}_l \triangleq \mathbf{y}(\tau)|_{\tau=lT_s}$. $\mathbf{R}_y$ can be estimated via

$$\hat{\mathbf{R}}_y = \frac{1}{L}\sum_{l=1}^{L}\mathbf{y}_l\mathbf{y}_l^H \quad (16)$$

B. PRE-WHITENING

The data model in Eq. (12) is non-redundant and more brief than that in Eq. (6). Nevertheless, the OGSBL framework based on Eq. (12) is still suffer from the underestimation of the noise variance. In this subsection, the pre-whitening processing is carried out to normalized the noise variance. Now, we define the error matrix

$$\Delta\mathbf{R}_y = \mathbf{R}_y - \hat{\mathbf{R}}_y \quad (17)$$

Therefore, we have

$$\begin{aligned}\mathrm{vec}\left(\Delta\mathbf{R}_y\right) &= \mathrm{vec}\left(\mathbf{F}^{-1/2}\mathbf{C}^H\left(\mathbf{R}_x - \hat{\mathbf{R}}_x\right)\mathbf{C}\mathbf{F}^{-1/2}\right) \\ &= \left(\left(\mathbf{F}^{-1/2}\mathbf{C}^H\right)^* \mathbf{F}^{-1/2}\mathbf{C}^H\right)\mathrm{vec}\left(\Delta\mathbf{R}_x\right) \\ &= \mathbf{Q}\mathrm{vec}\left(\Delta\mathbf{R}_x\right),\end{aligned} \quad (18)$$

where $\mathbf{R}_x$ and $\hat{\mathbf{R}}_x$ account for, respectively, the covariance matrix and its estimation of $\mathbf{x}(\tau)$, $\Delta\mathbf{R}_x = \mathbf{R}_x - \hat{\mathbf{R}}_x$, $\mathbf{Q} = \left(\mathbf{F}^{-1/2}\mathbf{C}^H\right)^* \mathbf{F}^{-1/2}\mathbf{C}^H$. It has been proven in the previous literature that $\mathrm{vec}\left(\Delta\mathbf{R}_x\right)$ satisfy an asymptotic Gaussian distribution [30], Since Gaussian distribution is also called Normal distribution, asymptotic Gaussian distribution is often denoted by AsN, thus

$$\mathrm{vec}\left(\Delta\mathbf{R}_x\right) \sim \mathrm{AsN}\left(\mathbf{0}, \mathbf{\Gamma}_x\right), \quad (19)$$

where $\mathbf{\Gamma}_x = 1/L\left(\mathbf{R}_x^T \otimes \mathbf{R}_x\right)$, and it can be approximated via $\hat{\mathbf{\Gamma}}_x = 1/L\left(\hat{\mathbf{R}}_x^T \otimes \hat{\mathbf{R}}_x\right)$. Accordingly, $\mathrm{vec}\left(\Delta\mathbf{R}_y\right)$ fulfills the following asymptotic Gaussian distribution

$$\mathrm{vec}\left(\Delta\mathbf{R}_y\right) \sim \mathrm{AsN}\left(\mathbf{0}, \mathbf{\Gamma}_y\right), \quad (20)$$

where $\mathbf{\Gamma}_y = \mathbf{Q}\hat{\mathbf{\Gamma}}_x\mathbf{Q}^H$. Denote $\mathbf{W} = \mathbf{R}_y^{-1/2}$, then the weighted covariance vector, $\sqrt{L}\left(\mathbf{W}^* \otimes \mathbf{W}\right)\mathrm{vec}\left(\Delta\mathbf{R}_y\right)$, fulfills the asymptotic standard normal distribution, i.e.,

$$\sqrt{L}\left(\mathbf{W}^* \otimes \mathbf{W}\right)\mathrm{vec}\left(\Delta\mathbf{R}_y\right) \sim \mathrm{AsN}\left(\mathbf{0}, \mathbf{I}\right) \quad (21)$$

Eq. (21) revels that the elements of $\sqrt{L}\mathbf{W}\Delta\mathbf{R}_y\mathbf{W}^H$ are independent and fulfill the standard normal distribution. As a result, we construct

$$\begin{aligned}\mathbf{Z} &= \sqrt{L}\mathbf{W}\left(\hat{\mathbf{R}}_y - \sigma^2\mathbf{I}\right)\mathbf{W}^H \\ &= \sqrt{L}\mathbf{W}\left(\mathbf{F}^{1/2}\mathbf{B}\mathbf{R}_s\mathbf{B}^H\mathbf{F}^{1/2} + \Delta\mathbf{R}_y\right)\mathbf{W}^H \\ &= \mathbf{G}\mathbf{B}\mathbf{D} + \mathbf{N},\end{aligned} \quad (22)$$

where $\mathbf{G} = \mathbf{W}\mathbf{F}^{1/2}$, $\mathbf{D} = \sqrt{L}\mathbf{R}_s\mathbf{B}^H\mathbf{F}^{1/2}\mathbf{W}^H$, $\mathbf{N} = \sqrt{L}\mathbf{W}\Delta\mathbf{R}_y\mathbf{W}^H$. Now we focus on $\mathbf{D}$, the covariance matrix of which is

$$\begin{aligned}\mathbf{\Gamma}_d &= L/M\mathbf{R}_s\mathbf{B}^H\mathbf{F}^{1/2}\mathbf{R}_y^{-1}\mathbf{F}^{1/2}\mathbf{B}\mathbf{R}_s^H \\ &= L/M\mathbf{R}_s\mathbf{B}^H\mathbf{F}^{1/2}\left(\mathbf{F}^{1/2}\mathbf{B}\mathbf{R}_s\mathbf{B}^H\mathbf{F}^{1/2} + \sigma^2\mathbf{I}\right)^{-1} \\ &\quad \times \mathbf{F}^{1/2}\mathbf{B}\mathbf{R}_s^H \\ &\approx L/M\mathbf{R}_s\mathbf{B}^H\left(\mathbf{B}\mathbf{R}_s\mathbf{B}^H\right)^\dagger\mathbf{B}\mathbf{R}_s^H \\ &= L/M\mathbf{R}_s,\end{aligned} \quad (23)$$

where the superscript $\dagger$ denotes the pseudo-inverse, the approximation is established since the noise can be omitted. Eq. (23) implies that each row of $\mathbf{D}$ are uncorrelated.

C. THE PROPOSED OGSBL

Actually, targets can be regarded to be sparse in the background. A fixed DOA grid $\varphi_1, \varphi_2, \cdots, \varphi_P$ $(K < N \ll P)$ is formed through the dispersion of possible directions. As a result, a dictionary matrix $\mathbf{\Psi} = [\mathbf{b}(\varphi_1), \mathbf{b}(\varphi_2), \cdots, \mathbf{b}(\varphi_P)]$ is obtained. It should be emphasized that there are K grids $\varphi_{\Omega_1}, \varphi_{\Omega_2}, \cdots, \varphi_{\Omega_K}$ that lying nearest to $\theta_1, \theta_2, \cdots, \theta_K$, $\{\Omega_k\}_{k=1}^K$ is the associate indexes in the set $\{1, 2, \cdots, P\}$. The off-grid DOA estimation of $\mathbf{Z}$ can be expressed as

$$\begin{aligned}\mathbf{Z} &= \mathbf{G}(\mathbf{\Psi} + \mathbf{\Phi})\mathbf{X} + \mathbf{N} \\ &= \mathbf{\Theta}\mathbf{X} + \mathbf{N},\end{aligned} \quad (24)$$

where $\mathbf{\Phi}$ is the perturbation to $\mathbf{\Psi}$, $\mathbf{\Theta} = \mathbf{G}(\mathbf{\Psi} + \mathbf{\Phi})$. In this paper, it is approximated by the first-order derivative of $\mathbf{b}\left(\varphi_p\right)$ with respective to φ_p $(p \in \{1, 2, \cdots, P\})$, i.e.,

$$\mathbf{\Phi} = \left[\alpha_1\frac{\partial\mathbf{b}(\varphi_1)}{\partial\varphi_1}, \alpha_2\frac{\partial\mathbf{b}(\varphi_2)}{\partial\varphi_2}, \cdots, \alpha_P\frac{\partial\mathbf{b}(\varphi_P)}{\partial\varphi_P}\right] \quad (25)$$

where α_p is the grid interval between the real DOA θ_k $(k \in \{1, 2, \cdots, K\})$ and its nearest grid φ_p, $\mathbf{X}$ is the sparse coefficient matrix that

$$\begin{cases}\alpha_p = \varphi_p - \theta_k, & \mathbf{X}_{p,\cdot} = \mathbf{D}_{k,\cdot} \quad if\ \varphi_p = \varphi_{\Omega_k} \\ \alpha_p = 0, & \mathbf{X}_{p,\cdot} = \mathbf{0}, \quad \text{otherwise}\end{cases} \quad (26)$$

where $\mathbf{X}_{p,\cdot}$ accounts for the p-th row of $\mathbf{X}$ and similar to others. In Eq. (24), $\mathbf{Z}$ and $\mathbf{\Psi}$ are known. Once the support of $\mathbf{X}$ and the intervals α_p $(p = 1, 2, \cdots, P)$ are obtained, the DOAs estimation problem is accomplished. Hence DOA estimation is linked to the following optimization problem

$$\arg\min_{\mathbf{X},\mathbf{\Phi}} \quad \|\mathbf{Z} - \mathbf{\Theta}\mathbf{X}\|_F^2 + \varepsilon f(\mathbf{X}), \quad (27)$$

where $f(\mathbf{X})$ is a penalty term that encourage sparsity, which various with different recovery strategies. In this paper, the OGSBL framework is utilized to solve the problem in Eq. (27).

As one of a statistical method, OGSBL is relay on the linear system and additive Gaussian noise model. Before the detailed derivation of the OGSBL algorithm, the following assumptions are given:

A.1: A hierarchical prior is assigned to $\mathbf{X}$. The rows of $\mathbf{X}$ are Gaussian process and share the same structural prior $\mathbf{H} \in \mathbb{C}^{P\times P}$. The density of $\mathbf{X}_{p,\cdot}$ $(p = 1, 2, \cdots, P)$ is given by

$$p\left(\mathbf{X}_{p,\cdot}; \gamma_p, \mathbf{H}\right) \sim \mathcal{N}\left(0, \gamma_p \mathbf{H}\right) \tag{28}$$

A.2: A Gamma hyper-prior is assigned to γ_p

$$p\left(\gamma_p\right) \sim \Gamma\left(\gamma_p \mid 1, \rho\right) \tag{29}$$

where ρ is an empirical coefficient. In this paper, we set $\rho \to 0.01$.

A.3: The grid interval α_p satisfies a uniform prior

$$p\left(\alpha_p\right) \sim \mathcal{U}\left([-1/2r, 1/2r]\right), \tag{30}$$

where r is the interval distance of the uniform grid $\varphi_1, \varphi_2, \cdots, \varphi_P$.

To explore the row sparse property of $\mathbf{X}$, the matrix model in Eq. (24) is formulated into vector version as

$$\mathbf{z} = \boldsymbol{\Upsilon}\mathbf{x} + \mathbf{n}, \tag{31}$$

where $\mathbf{z} = \mathrm{vec}\left(\mathbf{Z}^T\right)$, $\boldsymbol{\Upsilon} = \boldsymbol{\Theta} \otimes \mathbf{I}$, $\mathbf{s} = \mathrm{vec}\left(\mathbf{S}^T\right)$ and $\mathbf{n} = \mathrm{vec}\left(\mathbf{N}^T\right)$. According to **A.1**, the prior for $\mathbf{s}$ is given by

$$p\left(\mathbf{x}; \boldsymbol{\gamma}, \mathbf{H}\right) \sim \mathcal{N}\left(0, \boldsymbol{\Sigma}_0\right), \tag{32}$$

where $\boldsymbol{\gamma} = [\gamma_1, \gamma_2, \cdots, \gamma_P]^T$, $\boldsymbol{\Sigma}_0 = \boldsymbol{\Sigma} \otimes \mathbf{H}$ with $\boldsymbol{\Sigma} = \mathrm{diag}\left(\gamma_1, \gamma_2, \cdots, \gamma_P\right)$. Under the above assumptions, the Gaussian likelihood of $\mathbf{z}$ is

$$p\left(\mathbf{z} \mid \mathbf{x}; \boldsymbol{\Upsilon}\right) = \left(\frac{1}{\pi}\right)^{N^2} \exp\left\{- \parallel \mathbf{z} - \boldsymbol{\Upsilon}\boldsymbol{x} \parallel_2^2\right\} \tag{33}$$

According to the Bayes rule, the posterior density $p\left(\mathbf{x} \mid \mathbf{y}\right)$ is still Gaussian and can be expressed as

$$p\left(\mathbf{x} \mid \mathbf{z}; \boldsymbol{\Upsilon}, \boldsymbol{\gamma}\right) = \mathcal{N}\left(\boldsymbol{\mu}_x, \boldsymbol{\Sigma}_x\right), \tag{34}$$

with

$$\begin{cases} \boldsymbol{\mu}_x = \boldsymbol{\Sigma}_x \boldsymbol{\Upsilon}^H \mathbf{z} \\ \boldsymbol{\Sigma}_x = \left(\boldsymbol{\Upsilon}^H \boldsymbol{\Upsilon} + \boldsymbol{\Sigma}_0\right)^{-1} \end{cases} \tag{35}$$

Herein, the EM strategy is adopted to update the hyper-parameters $\boldsymbol{\Lambda} = \{\boldsymbol{\gamma}, \boldsymbol{\Phi}, \mathbf{H}\}$. To this end, we try to maximize $p\left(\mathbf{z}; \boldsymbol{\Lambda}\right)$, which is equal to minimize $-\ln p\left(\mathbf{z}; \boldsymbol{\Lambda}\right)$. The cost function of the EM strategy is

$$\begin{aligned} f\left(\boldsymbol{\Lambda}\right) &= \mathrm{E}_{\mathbf{x}|\mathbf{z};\boldsymbol{\Lambda}^{(\mathrm{old})}}\left[\ln p\left(\mathbf{z}, \mathbf{x}; \boldsymbol{\Lambda}\right)\right] \\ &= \mathrm{E}_{\mathbf{x}|\mathbf{z};\boldsymbol{\Lambda}^{(\mathrm{old})}}\left[\ln p\left(\mathbf{z} \mid \mathbf{x}; \boldsymbol{\Phi}\right)\right] \\ &\quad + \mathrm{E}_{\mathbf{x}|\mathbf{z};\boldsymbol{\Lambda}^{(\mathrm{old})}}\left[\ln p\left(\mathbf{x}; \boldsymbol{\gamma}, \mathbf{H}\right)\right], \end{aligned} \tag{36}$$

where $\boldsymbol{\Lambda}^{(\mathrm{old})}$ denote the parameter set in the last iteration.

To update one of the hyper-parameters, EM strategy treat all but one of the hyper-parameters as hidden variables. Therefore, it obtain the update rule by setting the derivative of $f\left(\boldsymbol{\Lambda}\right)$ with respect to the non-hidden variable to zero. Using this method, the learning rules for $\boldsymbol{\gamma}$ and $\mathbf{H}$ are

$$\boldsymbol{\gamma}_p = \frac{-N + \sqrt{N^2 + 4\rho \mathrm{Tr}\left[\mathbf{H}^{-1}\left(\boldsymbol{\Sigma}_x^p + \boldsymbol{\mu}_x^p\left(\boldsymbol{\mu}_x^p\right)^H\right)\right]}}{2\rho}, \tag{37}$$

and

$$\mathbf{H} = \left\{\sum_{p=1}^{P} \frac{\boldsymbol{\Sigma}_x^p + \boldsymbol{\mu}_x^p\left(\boldsymbol{\mu}_x^p\right)^H}{\gamma_p}\right\} / P, \tag{38}$$

where $\boldsymbol{\mu}_x^p$ denotes the p-th block (with size $M \times 1$). Since the updates of (35), (37) and (38) are alternate in high dimensional space, the iteration is computationally inefficient. In this paper, a fast version is derived by combining the MSBL idea. It calculate the mean matrix and covariance matrix in the original space as

$$\begin{cases} \tilde{\mathbf{X}} = \tilde{\boldsymbol{\Sigma}}\boldsymbol{\Theta}^H \mathbf{Z} \\ \tilde{\boldsymbol{\Sigma}} = \left(\boldsymbol{\Theta}^H \boldsymbol{\Theta} + \boldsymbol{\Sigma}\right)^{-1} \end{cases} \tag{39}$$

The learning rule for $\boldsymbol{\gamma}_p$ is derived as

$$\boldsymbol{\gamma}_p = \frac{-M + \sqrt{M^2 + 4\rho\left[\tilde{\boldsymbol{\Sigma}}_{p,p} + \tilde{\mathbf{X}}_{p\cdot}^* \mathbf{H}^{-1} \tilde{\mathbf{X}}_{p\cdot}^T\right]}}{2\rho}, \tag{40}$$

where $\tilde{\boldsymbol{\Sigma}}_{p,p}$ is the (p, p)-th entity of $\tilde{\boldsymbol{\Sigma}}$, and the learning rule for $\mathbf{H}$ is

$$\mathbf{H} = \frac{\tilde{\mathbf{H}} + \eta \mathbf{I}}{\|\tilde{\mathbf{H}} + \eta \mathbf{I}\|_F}, \tag{41}$$

with

$$\tilde{\mathbf{H}} = \sum_{p=1}^{P} \frac{\mathbf{X}_{p,\cdot}^T \mathbf{X}_{p,\cdot}^*}{\gamma_p}, \tag{42}$$

where η is a balance parameter that ensures $\mathbf{H}$ is positive define, and it set to $\eta = 2\|\tilde{\mathbf{H}}\|_F$ empirically. Finally, we concerning on $\boldsymbol{\Phi}$, which can be rewritten as

$$\boldsymbol{\Phi} = \mathbf{B}' \mathrm{diag}\left(\alpha_1, \alpha_2, \cdots, \alpha_P\right), \tag{43}$$

where $\mathbf{B}' = \left[\frac{\partial \mathbf{b}(\varphi_1)}{\partial \varphi_1}, \frac{\partial \mathbf{b}(\varphi_2)}{\partial \varphi_2}, \cdots, \frac{\partial \mathbf{b}(\varphi_P)}{\partial \varphi_P}\right]$, which is also known to us. Let $\mathbf{a} = [\alpha_1, \alpha_2, \cdots, \alpha_P]^T$, one can easy observe that the update of $\boldsymbol{\Phi}$ is rely on the update of $\mathbf{a}$. According to [27], $\mathbf{a}$ can be updated via

$$\arg \min_{\alpha_p \in [-1/2r, 1/2r]} \quad \mathbf{a}^T \mathbf{T} \mathbf{a} - 2\mathbf{v}^T \mathbf{a}, \tag{44}$$

with

$$\mathbf{T} = \mathrm{real}\left[\left(\tilde{\mathbf{B}}'^H \tilde{\mathbf{B}}'\right)^* \odot \left(\tilde{\boldsymbol{\Sigma}} + \sum_{m=1}^{M} \tilde{\mathbf{X}}_{\cdot,m} \tilde{\mathbf{X}}_{\cdot,m}^H\right)\right], \tag{45}$$

$$\mathbf{v} = \frac{1}{M}\text{real}\left[\sum_{m=1}^{M}\text{diag}\left(\tilde{\mathbf{X}}_{\cdot,m}^{*}\right)\left(\tilde{\mathbf{B}}'\right)^{H}\left(\mathbf{Z}_{\cdot,m} - \tilde{\mathbf{B}}\tilde{\mathbf{X}}_{\cdot,m}\right)\right] - \frac{1}{M}\text{real}\left[\text{diag}\left(\tilde{\mathbf{B}}'^{H}\tilde{\mathbf{B}}\tilde{\mathbf{\Sigma}}\right)\right], \quad (46)$$

where $\tilde{\mathbf{B}}' = \mathbf{G}\mathbf{B}'$ and $\tilde{\mathbf{B}} = \mathbf{G}\mathbf{B}$, $\tilde{\mathbf{X}}_{\cdot,m}$ denotes the m-th column of $\tilde{\mathbf{X}}$. The latter diag $(\cdot)$ in (46) returns a column vector. It is recommend by the author that if $\mathbf{T}$ is invertible, $\mathbf{a}$ is updated via

$$\mathbf{a} = \mathbf{T}^{-1}\mathbf{v} \quad (47)$$

Otherwise, it is updated entity-by-entity via computing

$$\tilde{\mathbf{a}}_p = \frac{\mathbf{v}_p - \left(\mathbf{T}_{p\cdot}\right)_{-p}\mathbf{a}_{-p}}{\mathbf{T}_{p,p}}, \quad (48)$$

and then calculate

$$\mathbf{a}_p = \begin{cases} \tilde{\mathbf{a}}_p, & \text{if } \tilde{\mathbf{a}}_p \in [-r/2, r/2] \\ -r/2, & \text{if } \tilde{\mathbf{a}}_p < -r/2 \\ r/2, & \text{otherwise}, \end{cases} \quad (49)$$

where $\mathbf{a}_{-p}$ returns a vector without the p-th element for a vector $\mathbf{a}$. The iteration will undergo before convergence, e.g., the iteration number reaches a pre-determined value, or the relative residual on $\boldsymbol{\gamma}$ is smaller than a given threshold. For more algorithmic details, the reader is recommend to refer [27].

Now we have achieved the proposal of the RC-OGSBL estimator. To help the reader to understand the proposed estimator, we list the main steps of RC-OGSBL as follows:

Step 1: Construct $\mathbf{C}$ according to (8), and get the non-redundant measurement $\mathbf{y}(\tau)$ via (12);

Step 2: Estimate the covariance matrix $\mathbf{R}_y$ via (16), and obtain $\mathbf{Z}$ via (22);

Step 3: Form $\mathbf{B}$ and $\mathbf{B}'$. Initial $\mathbf{a}$ $\mathbf{H}$ and $\boldsymbol{\gamma}$;

Step 4: Update the mean $\tilde{\mathbf{X}}$ and the covariance matrix $\tilde{\mathbf{\Gamma}}$ via (39);

Step 5: Update $\boldsymbol{\gamma}_p$ and $\mathbf{H}$ via (40) and (41), respectively;

Step 6: Refine the grid via (47) or (49);

Step 7: Repeat **step.4** to **step.6** until algorithm convergence.

IV. RELATED REMARKS

Remark 1: It is obvious that the proposed algorithm has blind character, while the MUSIC algorithm requires the prior information of the target number.

Remark 2: It should be pointed out that the proposed RC-OGSBL algorithm is only suitable for monostatic MIMO radar system with ULA geometry, otherwise it will fail to work.

Remark 3: The graphical model of the OGSBL framework in [27] and the graphical model of the improved OGSBL algorithm are illustrated in Figure 2 and Figure 3, respectively. In Figure 2 and Figure 3, the circles denote the parameters or signals, rectangles account for the hyper-parameters.

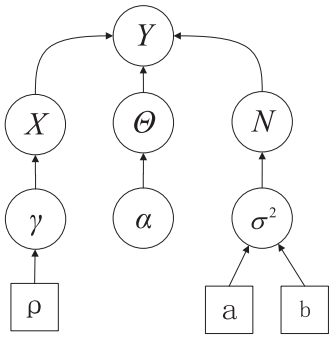


FIGURE 2. Graphical model of the OGSBL framework in [27].

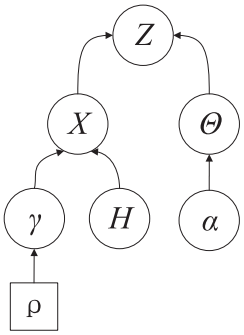


FIGURE 3. Graphical model of the improved OGSBL algorithm.

It is shown that the OGSBL in [27] requires necessary learning for the noise, but this process is neglected in the proposed algorithm. Besides, a blcok parameter $\mathbf{H}$ is given in our algorithm, which is helpful in accelerating convergence.

V. SIMULATION RESULTS

To verify the effectiveness of the improved OGSBL estimator, numerical simulations have been carried out. In this section, we consider a monostatic MIMO radar system, which is configured with M transmit antennas and N receive antennas, both of which are assumed to be ULAs with half-wavelength spacing. Suppose $K = 3$ uncorrelated sources, the reflection coefficients fulfill the Swerling II model, and L snapshots are collected. In the simulation, the grid interval is Δ, and the signal-to-noise ratio (SNR) is defined as $\text{SNR} = 10\log_{10}\|\mathbf{x}(\tau) - \mathbf{e}(\tau)\|^2 / \|\mathbf{e}(\tau)\|^2$ [dB], where $\mathbf{x}(\tau)$ and $\mathbf{e}(\tau)$ are the signal in (6). In *Example* 2-*Example* 5, $K = 3$ targets are considered with DOAs are set to 31.2°, 60.3° and 120.5°, respectively. The simulation is carried out on a HP Z840 workstation with two Intel(R) Xeon(R) E5-2650 v4 2.20 GHz processors, and 128 GB of RAM.

Example 1: We compare the spectrum of the proposed algorithm with RC-MUSIC in [34], where $M = 6$, $N = 8$, $L = 200$, $\Delta = 4°$ and $\text{SNR} = -10\text{dB}$. Figure 4 and Figure 5 illustrate the results with $K = 2$ and $K = 4$, respectively. It is seen that the although both algorithms work correctly, the proposed estimator provides much lower sidelobe then the MUSIC algorithm. Moreover, as shown in the small rectangles, the peaks of the proposed estimator is more closer to the

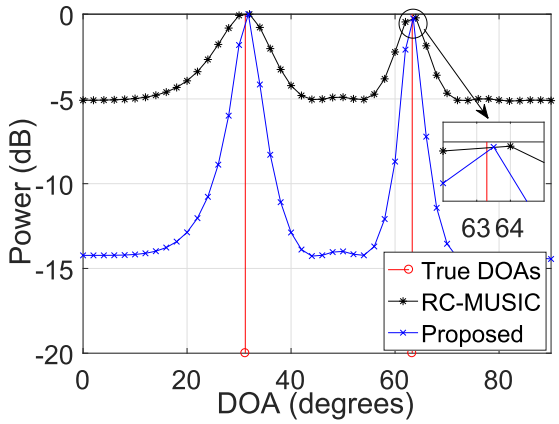


FIGURE 4. Spectrum comparison with $K = 2$ targets.

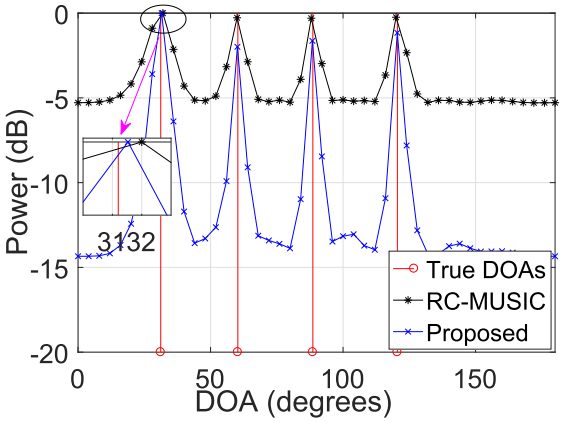


FIGURE 5. Spectrum comparison with $K = 4$ targets.

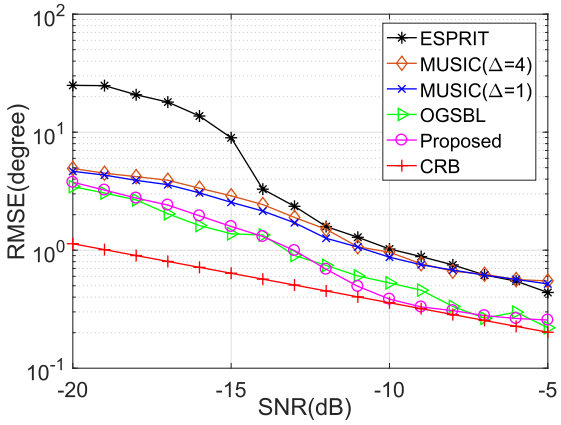


FIGURE 6. RMSE comparison of various method versus SNR.

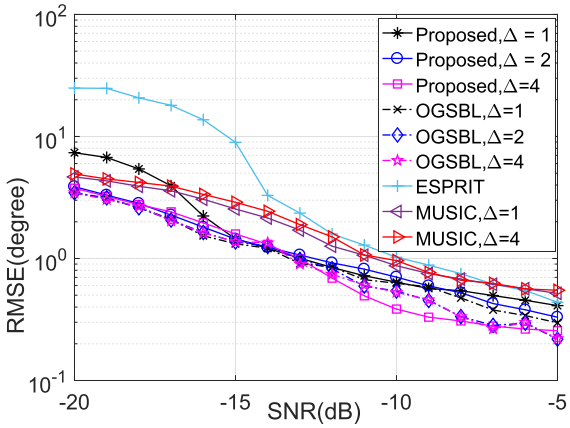


FIGURE 7. RMSE versus SNR with different Δ.

real DOAs, since the grid in the proposed algorithm can be updated adaptively, while it is fixed in MUSIC algorithm.

Example 2: We exam the estimation performance of the proposed estimator in terms of root mean square error (RMSE), where $M = 6$, $N = 8$, $L = 200$ and $\Delta = 4°$ are considered. The RMSE is defined as

$$\text{RMSE} = \frac{1}{K}\sum_{k=1}^{K}\sqrt{\frac{1}{500}\sum_{i=1}^{500}\left(\hat{\theta}_{i,k} - \theta_k\right)^2}, \tag{50}$$

where $\hat{\theta}_{i,k}$ and $\hat{\varphi}_{i,k}$ are the estimates of θ_k and φ_k in the i th trial. For comparison, the performance of the RC-ESPRIT [33] (marked with ESPRIT), RC-MUSIC in [34] (marked with MUSIC), OGSBL [27] as well as the Cramer-Rao bound (CRB) are added. All the curves are based on 500 Monte Carlo trials. One can observe from Figure 6 that RC-MUSIC algorithms provide better RMSE than RC-ESPRIT, as more DOF has been exploited in RC-MUSIC. Besides, RMSE performance of RC-MUSIC is improved with increasing SNR, and smaller search interval Δ results in more accurate DOA estimation performance. Obviously, both OGSBL and the proposed estimator have very close RMSE performance, and both of which provider much lower RMSE than RC-MUSIC. Their superior performance benefit from the fact that the grid of both estimators can be updated during the iteration, while the grid is fixed in MUSIC.

Example 3: We compared the estimation performance of OGSBL, RC-MUSIC and the proposed estimator with various Δ, where $M = 6$, $N = 8$ and $L = 200$ are set. Two meters are adopted herein, one is the RMSE and the other is the average running time. Also, the performance of RC-RSPRIT has been added. The results are given in Figure 7 and Figure 8. It is obvious that RC-MUSIC and RC-ESPRIT need less running time, as they does not involve iterative calculation. However, they provide higher RMSE than the Bayesian estimators (OGSBL and the proposed estimator). Moreover, smaller Δ leads to heavier computational load, as shown in Figure 8. Besides, the complexity of the proposed estimator is improved by one order of magnitude in contrast to the OGSBL algorithm. This improvement benefit from two aspects. On the one hand, the RC technique in the proposed estimator helps to reduce the dimension of the measurement without hurt the DOF of MIMO radar. On the other hand, the pre-whitening process in the proposed estimator makes the noise variance to be determined, thus reduces the parameter number in the iteration. It should be noticed that the proposed estimator share very close RMSE with the OGSBL

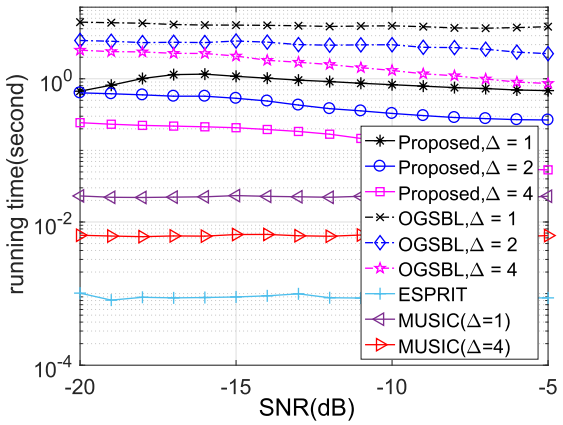


FIGURE 8. Average running time versus SNR with different Δ.

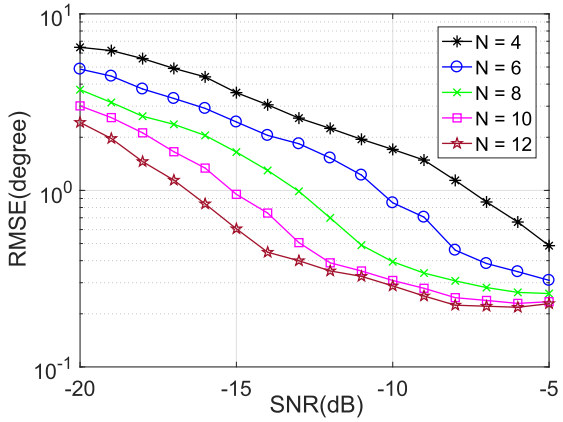


FIGURE 9. RMSE versus SNR with different *N*.

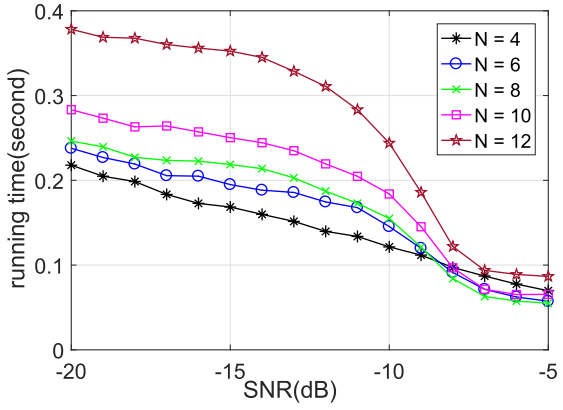


FIGURE 10. Average running time versus SNR with different *N*.

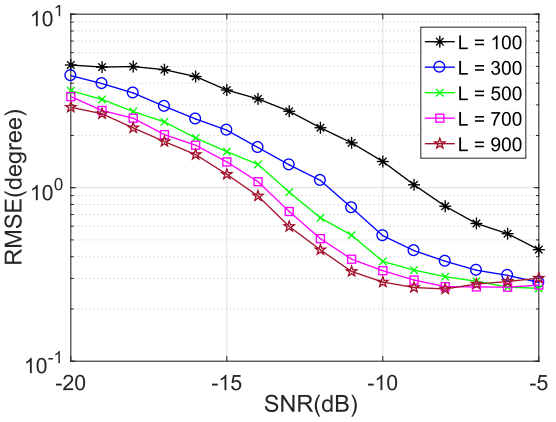


FIGURE 11. RMSE versus SNR with different *L*.

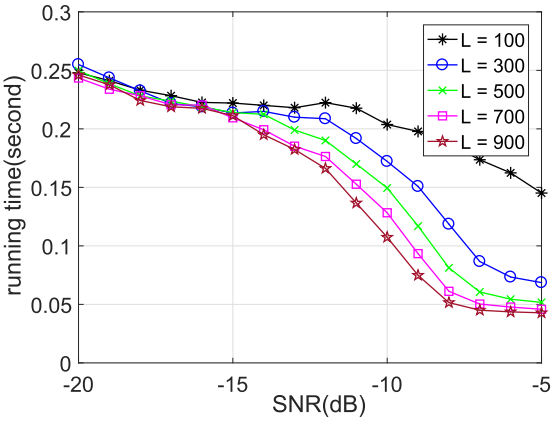


FIGURE 12. Average running time versus SNR with different *L*.

approach at low SNR regions. However, it may perform worse than OGSBL with high SNR, as depicted in Figure 7, since the formula (23) is approximate established.

Example 4: We repeat the performance test with various N, where $M = 6$, $L = 200$ and $\Delta = 4$. Figure 9 and Figure 10 present the performance curves. It is shown in Figure 9 that large N bring better DOA estimation accuracy, since larger N means the effective aperture of MIMO radar is larger. Similar to *Example* 3, more running time is required when N is increasing, but the necessary running time corresponding to different N is very close when SNR>-9dB.

Example 5: We test the estimation performance in terms of various L, where $M = 6$, $N = 8$ and $\Delta = 4$ are considered. The RMSE performance and the average running time performance are depicted in Figure 11 and Figure 12, respectively. It is displays that larger L results in fewer calculation, thus less running time is required. Furthermore, it is shown that more snapshot L leads to better RMSE performance, as the noise model is more accurate with larger L.

VI. CONCLUSION

In this paper, we revisit the DOA estimation problem in colocated MIMO radar with ULA geometries. A reduced-complexity OGSBL framework is presented. The proposed estimator can be summarized into three main steps. Firstly, it removes the redundancy of the array measurement via reduced-complexity transformation. Thereafter, a covariance matrix model is established by pre-whitening the non-redundant signal, in which the noise is white with variance is 1. Finally, the DOA estimation problem is linked to OGSBL with less hyper-parameters. The RC technique and the pre-whitening process reduce the redundancy of colocated MIMO radar without hurt the visual aperture, and they can

help to achieve an off-grid Bayesian model with less variables, which enable the proposed estimator more efficient than the OGSBL algorithm. The performance of the proposed algorithm is verified via numerical simulations.

REFERENCES

[1] D. Ravi, C. Wong, F. Deligianni, M. Berthelot, J. Andreu-Perez, B. Lo, and G.-Z. Yang, "Deep learning for health informatics," *IEEE J. Biomed. Health Inform.*, vol. 21, no. 1, pp. 4–21, Jan. 2017.

[2] H. Wang, L. Wan, M. Dong, K. Ota, and X. Wang, "Assistant vehicle localization based on three collaborative base stations via SBL-based robust DOA estimation," *IEEE Internet Things J.*, vol. 6 no. 3, pp. 5766–5777, Jun. 2019.

[3] H. Huang, J. Yang, H. Huang, Y. Song, and G. Gui, "Deep learning for super-resolution channel estimation and doa estimation based massive MIMO system," *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8549–8560, Sep. 2018.

[4] H. Huang, W. Xia, J. Xiong, J. Yang, G. Zheng, and X. Zhu, "Unsupervised learning-based fast beamforming design for downlink MIMO," *IEEE Access*, vol. 7, pp. 7599–7605, 2019.

[5] H. Huang, Y. Song, J. Yang, G. Gui, and F. Adachi, "Deep-learning-based millimeter-wave massive MIMO for hybrid precoding," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 3027–3032, Mar. 2019.

[6] Y. Wang, M. Liu, J. Yang, and G. Gui, "Data-driven deep learning for automatic modulation recognition in cognitive radios," *IEEE Trans. Veh. Technol.*, vol. 68, no. 4, pp. 4074–4077, Apr. 2019.

[7] G. Gui, H. Huang, Y. Song, and H. Sari, "Deep learning for an effective nonorthogonal multiple access scheme," *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8440–8450, Sep. 2018.

[8] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, pp. 2481–2495, Dec. 2017.

[9] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 4, pp. 834–848, Apr. 2018.

[10] 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.

[11] F. Wen, "Computationally efficient DOA estimation algorithm for MIMO radar with imperfect waveforms," *IEEE Commun. Lett.*, vol. 23, no. 6, pp. 1037–1040, Jun. 2019.

[12] 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.

[13] H. Chen, W. Wang, and W. Liu, "Joint DOA, range, and polarization estimation for rectilinear sources with a cold array," *IEEE Wireless Commun. Lett.*, to be published. doi: 10.1109/LWC.2019.2919542.

[14] 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.

[15] 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.

[16] 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.

[17] F. Wen, Z. Zhang, G. Zhang, Y. Zhang, X. Wang, and X. Zhangc, "A tensor-based covariance differencing method for direction estimation in bistatic MIMO radar with unknown spatial colored noise," *IEEE Access*, vol. 5, pp. 18451–18458, Sep. 2017.

[18] 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, Sep. 2017.

[19] F. Wen, Z. Zhang, and G. Zhang, "Joint DOD and DOA estimation for bistatic MIMO radar: A covariance trilinear decomposition perspective," *IEEE Access*, vol. 7, no. 1, pp. 53273–53283, Apr. 2019.

[20] J. F. Li, X. F. Zhang, and X. Gao, "A joint scheme for angle and array gain-phase error estimation in bistatic MIMO radar," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 6, pp. 1478–1482, Nov. 2013.

[21] J. Li and X. Zhang, "A method for joint angle and array gain-phase error estimation in bistatic multiple-input multiple-output non-linear arrays," *IET Signal Process.*, vol. 8, no. 2, pp. 131–137, Apr. 2014.

[22] F. Wen, X. Xiong, and Z. Zhang, "Angle and mutual coupling estimation in bistatic MIMO radar based on PARAFAC decomposition," *Digit. Signal Process.*, vol. 65, pp. 1–10, Jun. 2017.

[23] F. Wen, Z. Zhang, K. Wang, G. Sheng, and G. Zhang, "Angle estimation and mutual coupling self-calibration for ULA-based bistatic MIMO radar," *Signal Process.*, vol. 144, pp. 61–67, Mar. 2018.

[24] F. Wen, Z. Zhang, and X. Zhang, "CRBs for direction-of-departure and direction-of-arrival estimation in collocated MIMO radar in the presence of unknown spatially coloured noise," *IET Radar, Sonar Navigat.*, vol. 13, no. 4, pp. 530–537, Apr. 2019.

[25] F. Wen, C. Mao, and G. Zhang, "Direction finding in MIMO radar with large antenna arrays and nonorthogonal waveforms," *Digit. Signal Process.*, to be published. doi: 10.1016/j.dsp.2019.06.008.

[26] H. Zhu, G. Leus, and G. B. Giannakis, "Sparsity-cognizant total least-squares for perturbed compressive sampling," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2002–2016, May 2011.

[27] Z. Yang, L. Xie, and C. Zhang, "Off-grid direction of arrival estimation using sparse Bayesian inference," *IEEE Trans. Signal Process.*, vol. 61, no. 1, pp. 38–43, Jan. 2013.

[28] H. Zamani, H. Zayyani, and F. Marvasti, "An iterative dictionary learning-based algorithm for DOA estimation," *IEEE Commun. Lett.*, vol. 20, no. 9, pp. 1784–1787, Sep. 2016.

[29] J. Dai, X. Bao, W. Xu, and C. Chang, "Root sparse Bayesian learning for off-grid DOA estimation," *IEEE Signal Process. Lett.*, vol. 24, no. 1, pp. 46–50, Jan. 2017.

[30] Y. Zhang, Z. Ye, X. Xu, and N. Hu, "Off-grid DOA estimation using array covariance matrix and block-sparse Bayesian learning," *Signal Process.*, vol. 98, no. 1, pp. 197–201, May 2014.

[31] P. Chen, Z. Cao, Z. Chen, and X. Wang, "Off-grid DOA estimation using sparse Bayesian learning in MIMO radar with unknown mutual coupling," *IEEE Trans. Signal Process.*, vol. 67, no. 1, pp. 208–220, Jan. 2019.

[32] B. Liao, "Fast angle estimation for MIMO radar with nonorthogonal waveforms," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 54, no. 4, pp. 2091–2096, Aug. 2018.

[33] X. Zhang and D. Xu, "Low-complexity ESPRIT-based DOA estimation for colocated MIMO radar using reduced-dimension transformation," *Electron. Lett.*, vol. 47, no. 4, pp. 283–284, Feb. 2011.

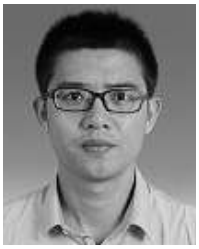
[34] Y. Zhao, P. Shui, and H. Liu, "Computationally efficient DOA estimation for MIMO radar," in *Proc. 2nd Int. Congr. Image Signal Process.*, Oct. 2009, pp. 1–3.

[35] Y. Zhang, G. Zhang, and X. Wang, "Computationally efficient DOA estimation for monostatic MIMO radar based on covariance matrix reconstruction," *Electron. Lett.*, vol. 53, no. 2, pp. 111–113, Jan. 2017.

[36] F. Wen, D. Huang, K. Wang, and L. Zhang, "DOA estimation for monostatic MIMO radar using enhanced sparse Bayesian learning," *J. Eng.*, vol. 2018, no. 5, pp. 268–273, May 2018.

TINGTING LIU was born in Hubei, China, in 1990. She received the B.S. degree in marketing from GanNan Normal University, Ganzhou, China, in 2012, and the M.S. degree in business management, Nanjing University of Finance and Economics, China, in 2016. Since January 2017, she has been with the School of Physical, Yangtze University, China, where she is currently an Associate Professor. Her research interests include gray system, sensor array, and optimization.

FANGQING WEN was born in Hubei, China, in 1988. He received the B.S. degree in electronic engineering from the Hubei University of Automotive Technology, Shiyan, China, in 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 with NUAA, in 2016. From October 2015 to April 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 (MIMO) radar, array signal processing, and compressive sensing. He is a member of the Chinese Institute of Electronics (CIE). He is also an Associate Editor of the *Electronics Letters* journal.



LEI ZHANG was born in Hubei, China, in 1987. He received the B.S. degree in information and computing science from the China University of Geosciences, Wuhan, China, in 2010, the master's degree from the College of Information and Communication Engineering, Harbin Engineering University (HEU), China, in 2012, and the Ph.D. degree from HEU, in 2016. He is currently a Lecturer with the Electronic and Information School, Yangtze University, China. His research interests include intelligent information processing and image processing.

KE WANG was born in Hubei, China, in 1988. He received the B.S. degree in automation engineering from the Hubei University of Technology, Wuhan, China, in 2010, the master's degree from the College of Electronic Information and Control Engineering, Beijing University of Technology, China, in 2016, and the Ph.D. degree from the Beijing University of Technology, in 2016. Since 2016, he has been with the Electronic and Information School, Yangtze University, China, where he is currently a Lecturer. His research interests include mobile robot and computer vision.