A Fast Gridless Sparse Method for Robust DOA Estimation in the Presence of Gain-phase Errors

Wencan Peng [1], Chenjiang Guo [1], Min Wang [2], Yuteng Gao [1], and Xia Zhao [1]

[1] School of Electronics and Information, Northwestern Polytechnical University, Xi'an, 710129, China
wcanpeng@163.com, ciguo@nwpu.edu.cn, ytgao2002@163.com, zhaoxia908@126.com

[2] National Lab of Radar Signal Processing, Xidian University, Xi'an, 710071, China
wangmin@xidian.edu.cn

Abstract – A new gridless sparse method (GLSM) is proposed to estimate the direction of arrival (DOA) and gain-phase errors simultaneously for a uniform linear array (ULA). We convert angular space to frequency space and establish a data model in the frequency domain. First, the cost function based on the covariance fitting criterion is transformed into a semidefinite programming (SDP) problem to estimate DOA and noise variance without previous calibration information. Second, gain errors are calculated by the estimated noise variance and the covariance matrix. Third, phase errors are obtained by decomposition of the covariance matrix, which has been pre-processed by a space smoothing technique. Finally, DOA estimation is improved further after the array errors are fully calibrated. Compared with traditional methods, the proposed method is robust to correlations of signal sources, and parameters are estimated without joint iteration. Moreover, there is no need for discrete grid points in the angular space, which results in grid mismatches and computation loads, so the proposed method is more accurate and faster. Simulation results verify the effectiveness of the proposed method.

Index Terms – Direction of arrival, gain and phase errors, gridless sparse method, semidefinite programming.

I. INTRODUCTION

The direction of arrival (DOA) estimation is one of the most important topics in array signal processing. Many high resolution DOA algorithms have been proposed, such as multiple signal classification (MUSIC) [1], estimation of signal parameters via rotational invariance techniques (ESPRIT) [2], maximum likelihood [3] and Capon [4]. Recently, DOA estimation performance has been further improved by sparse reconstruction algorithms. To study the sparsity of DOAs, a spatial domain is usually discretized to create a dictionary to sparsely represent the DOAs [5-7]. The l_1 -based singular value decomposition (L1_SVD) algorithm [5], based on an overcomplete basis, handles the DOA estimation problem through SVD and a second-order cone (SOC) framework. The method in [6] based on the covariance fitting criteria estimates the source power and the corresponding DOAs by an iteration process. However, those on-grid methods assume that the true DOAs are located on a fixed set of grid points, so grid mismatch is inevitable. The off-grid methods [8-10] and gridless methods [11-12] are used to solve the grid mismatch problem. Additionally, the latter completely overcome this problem by operating in the continuous domain. A novel off-grid method [10] is proposed via successive nonconvex sparsity approximation penalties on the sparse signals and jointly estimates the sparse signals and grid offset parameters. The method proposed in [11] is a grid-free sparse method. DOAs and noise variance are solved by semidefinite programming (SDP) and convex optimization. However, all those subspace and sparse methods have the common assumption of ideal array manifolds. In practice, the array manifold is often affected by unknown array perturbations, such as gain and phase uncertainties, mutual coupling and sensor location errors [13-15]. Our paper focuses on the gain and phase errors of antenna arrays.

To deal with array gain-phase perturbations, array calibration algorithms are usually classified into two groups. The first group is called active calibration algorithms [16-18]. These methods require that all the DOAs are precisely known. With the use of calibration sources, the array can be calibrated well. However, it is hard to accurately determine the DOAs of the calibration sources in practice [19]. The second group is called self-calibration algorithms, which are more desirable. They can estimate both the array perturbations and DOA without knowing the exact locations of calibration sources. In [20], the method was proposed by Weiss and Friedlander and called the WF method for convenience. DOAs and gain-phase errors are estimated

Submitted On: March 26, 2019
Accepted On: August 11, 2019

by an alternative iteration method in the case of uncorrelated sources. However, convergence to the global optimum cannot be guaranteed when the phase error is large. In [21] and [22], non-iterative methods based on the WF method for phase errors are proposed. These methods have the benefit that the term of gain-phase errors is eliminated from the array manifold by the Hadamard product, and DOAs and gain-phase errors are solved independently. However, they have the disadvantages that the computation is complicated, the case of coherent signals is not considered, and they are not applicable to a uniform linear array (ULA). The method in [23] is a sparse-based method to achieve a joint estimation of DOA and array perturbations, which is based on a spatial dictionary, so it encounters the following problems: increased computational loads and errors caused by discretization. In [24], an on-grid method (ONGM) based on the covariance fitting criteria is proposed to estimate the DOA and gain-phase error simultaneously by alternate iteration. It performs well in cases of uncorrelated sources and coherent sources. However, the main drawbacks of ONGM are converging to suboptimal solutions in the case of large phase errors and high computational complexity.

Inspired by [24], we propose a gridless sparse method (GLSM) to deal with the DOA estimation problem of gain-phase errors. The covariance fitting criteria have sound statistical motivation and are robust to coherent signal sources [6, 25]. Moreover, according to [26], covariance fitting is related to the l_1 _norm minimization, which is robust to correlation. We consider the same covariance fitting criteria as the ONGM method, but the covariance fitting problem is transformed into an SDP problem and solved by SDPT3 or other SDP solvers [27-28] for estimating DOA and noise variance in the frequency domain. Then, the gain perturbations are estimated by using the covariance matrix and the estimated noise covariance, and phase errors are estimated by a common method with estimated DOAs. Finally, we use SDP solvers again to further improve DOA estimation after the gain-phase errors are compensated. Compared with the ONGM method [24], the proposed method overcomes some drawbacks. First, no discretization scheme is adopted, which avoids grid mismatch. Second, parameters are estimated with no iteration, so it does not fall into local optimality in the case of large phase errors and has better estimation performance. Third, the proposed method does not depend on the grid size and the number of iterations; thus, it can be faster than the ONGM method when a dense sampling grid is used in ONGM for high estimation precision. Simulation results show the effectiveness of the proposed method.

II. DATA MODEL

Consider that K narrowband far-field signals impinge on M isotropic sensors from different incident angles $\boldsymbol{\theta}=[\theta_1,\ldots,\theta_K]^{\mathrm{T}}$. The array spacing is half-wavelength and N is the number of snapshots. We convert the angular domain to the frequency domain to avoid discretizing the spatial angles.

We denote $f_k=(\sin(\theta_k)+1)/2\in[0,1),\ k=1,\ldots K,$ and the relation $\boldsymbol{\theta}\leftrightarrow\boldsymbol{f}$ is one-to-one. We assume that $K\le M-1$ in the following discussion. The steering matrix can thus be represented as $\boldsymbol{A}(\boldsymbol{f})=[\boldsymbol{a}(f_1),\ldots,\boldsymbol{a}(f_K)]$, with $\boldsymbol{a}(f_k)=[1,\mathrm{e}^{\mathrm{j}2\pi f_k},\ldots,\mathrm{e}^{\mathrm{j}2(M-1)\pi f_k}]^{\mathrm{T}}$. f_k is the frequency of the uniformly sampled complex sinusoid $\boldsymbol{a}(f_k)$, so $\boldsymbol{f}=[f_1,\ldots,f_K]^{\mathrm{T}}$ is the frequency parameter. Considering the gain-phase errors, the received signals at the ULA along the x-axis can be expressed as:

$$\boldsymbol{Y}=\boldsymbol{\Gamma}\boldsymbol{A}(\boldsymbol{f})\boldsymbol{S}+\boldsymbol{N}, \tag{1}$$

where $\boldsymbol{\Gamma}=\mathrm{diag}([\gamma_1,\gamma_2\ldots,\gamma_M]^{\mathrm{T}})$ denotes the gain-phase error diagonal matrix. Define $\gamma_m\triangleq\alpha_m e^{j\phi_m}$, and α_m and ϕ_m denote the gain error and phase error, respectively, of the mth sensor. In addition, we assume that $\alpha_1=1$, $\phi_1=0$, the noise $\boldsymbol{N}$ is additive white Gaussian noise and that the source signal $\boldsymbol{S}$ is independent of the noise.

The covariance matrix $\boldsymbol{R}$ can be written as:

$$\boldsymbol{R}(\boldsymbol{f},\boldsymbol{p},\boldsymbol{\sigma},\boldsymbol{\Gamma})=\boldsymbol{\Gamma}\boldsymbol{A}(\boldsymbol{f})\mathrm{diag}(\boldsymbol{p})\boldsymbol{A}^{\mathrm{H}}(\boldsymbol{f})\boldsymbol{\Gamma}^{\mathrm{H}}+\mathrm{diag}(\boldsymbol{\sigma}), \tag{2}$$

where $\boldsymbol{p}=[p_1,\ldots,p_K]^{\mathrm{T}}$ is the source power and $\boldsymbol{\sigma}=[\sigma^2,\ldots,\sigma^2]^{\mathrm{T}}_{M\times1}$ denotes the noise variance.

For simplicity, note,

$$\boldsymbol{C}(\boldsymbol{f},\boldsymbol{p})=\boldsymbol{A}(\boldsymbol{f})\mathrm{diag}(\boldsymbol{p})\boldsymbol{A}^{\mathrm{H}}(\boldsymbol{f}). \tag{3}$$

The (m,l)th element of $\boldsymbol{C}$ is:

$$C_{\mathrm{m},l}=\sum_{k=1}^{K}p_k^2 a_m(f_k)a_l^*(f_k)=\sum_{k=1}^{K}p_k^2 e^{j2\pi(\mathrm{m}-l)f_k}. \tag{4}$$

It is easy to see that $\boldsymbol{C}\ge0$ and $\boldsymbol{C}$ is a (Hermitian) Toeplitz matrix.

The sample covariance matrix can be expressed as:

$$\hat{\boldsymbol{R}}=\boldsymbol{Y}\boldsymbol{Y}^H/N. \tag{5}$$

III. THE PROPOSED METHOD

A. Robust DOA estimation

We consider the following cost function based on covariance fitting criterion [6, 25]:

$$g(\boldsymbol{f},\boldsymbol{p},\boldsymbol{\sigma},\boldsymbol{\Gamma})=\left\|\boldsymbol{R}^{-1/2}(\hat{\boldsymbol{R}}-\boldsymbol{R})\hat{\boldsymbol{R}}^{-1/2}\right\|_{\mathrm{F}}^2, \tag{6}$$

where $\|\bullet\|_{\mathrm{F}}$ represents the Frobenius norm for matrices and the l_2 _norm for vectors. According to [11], the minimization of the criterion in (6) is a large-snapshot realization of the ML estimator. A simple calculation is as follows:

$$\begin{aligned} g(\boldsymbol{f}, \boldsymbol{p}, \boldsymbol{\sigma}, \boldsymbol{\Gamma}) &= \mathrm{tr}[\boldsymbol{R}^{-1/2}(\hat{\boldsymbol{R}}-\boldsymbol{R})\hat{\boldsymbol{R}}^{-1}(\hat{\boldsymbol{R}}-\boldsymbol{R})\boldsymbol{R}^{-1/2}] \\ &= \mathrm{tr}[(\boldsymbol{R}^{-1}\hat{\boldsymbol{R}}-\boldsymbol{I})(\boldsymbol{I}-\hat{\boldsymbol{R}}^{-1}\boldsymbol{R})] \\ &= \mathrm{tr}(\boldsymbol{R}^{-1}\hat{\boldsymbol{R}})+\mathrm{tr}(\hat{\boldsymbol{R}}^{-1}\boldsymbol{R})-2M, \end{aligned} \tag{7}$$

where $\mathrm{tr}(\bullet)$ denotes the trace of matrices. In the presence of noise $\boldsymbol{R}^{-1}$ exists while $\hat{\boldsymbol{R}}^{-1}$ exists under the condition $N \geq M$. $\boldsymbol{I}$ denotes the identity matrix. The unknown parameters $\boldsymbol{f}$, $\boldsymbol{p}$, $\boldsymbol{\sigma}$ and $\boldsymbol{\Gamma}$ are nonlinear in relation to $\boldsymbol{R}$, so it is challenging to minimize g. To calculate those parameters in (7), it is necessary to reparameterize $\boldsymbol{R}$ for reducing the number of parameters. According to (3), $\boldsymbol{C}$ is a (Hermitian) Toeplitz matrix, so it can be expressed as $\boldsymbol{C}=T(\boldsymbol{u})$ depending on the first row $\boldsymbol{u}$. Substitute $\boldsymbol{C}=T(\boldsymbol{u})$ into (2) to obtain:

$$\boldsymbol{R}(\boldsymbol{u},\boldsymbol{\sigma},\boldsymbol{\Gamma}) = \boldsymbol{\Gamma}T(\boldsymbol{u})\boldsymbol{\Gamma}^{\mathrm{H}} + \mathrm{diag}(\boldsymbol{\sigma}). \tag{8}$$

According to [6], the cost function in (6) is convex and has a unique global minimum. Problem (7) can thus be formulated as the following SDP problem:

$$\begin{aligned} &\min_{\boldsymbol{\Gamma},\boldsymbol{u},\{\boldsymbol{\sigma}\succeq 0\}} \ \mathrm{tr}\left(\boldsymbol{R}^{-1}\hat{\boldsymbol{R}}\right)+\mathrm{tr}\left(\hat{\boldsymbol{R}}^{-1}\boldsymbol{R}\right) \\ &\text{subject to } T(\boldsymbol{u}) \geq 0, \end{aligned} \tag{9}$$

where $\boldsymbol{\sigma} \succeq \mathbf{0}$ represents that $\sigma_m \geq 0$ for all m, then we show the following equivalence:

$$\begin{aligned} (9) \Leftrightarrow &\min_{\boldsymbol{\Gamma},\boldsymbol{u},\{\boldsymbol{\sigma}\succeq 0\}} \ \mathrm{tr}\left(\hat{\boldsymbol{R}}^{\frac{1}{2}}\boldsymbol{R}^{-1}\hat{\boldsymbol{R}}^{\frac{1}{2}}\right)+\mathrm{tr}\left(\hat{\boldsymbol{R}}^{-1}\boldsymbol{R}\right) \\ &\text{subject to } T(\boldsymbol{u}) \geq 0. \end{aligned} \tag{10}$$

The parameters of $\boldsymbol{R}^{-1}$ in (10) are in the denominator, which does not satisfy the condition of CVX, so we introduce parameter $\boldsymbol{Z}$, satisfying the constraint condition $\boldsymbol{Z} \geq \hat{\boldsymbol{R}}^{1/2}\boldsymbol{R}^{-1}\hat{\boldsymbol{R}}^{1/2}$,

$$\begin{aligned} &\min_{\boldsymbol{Z},\boldsymbol{\Gamma},\boldsymbol{u},\{\boldsymbol{\sigma}\succeq 0\}} \ \mathrm{tr}\left(\boldsymbol{Z}\right)+\mathrm{tr}\left(\hat{\boldsymbol{R}}^{-1}\boldsymbol{R}\right) \\ &\quad \text{subject to } T(\boldsymbol{u}) \geq 0 \text{ and } \boldsymbol{Z} \geq \hat{\boldsymbol{R}}^{\frac{1}{2}}\boldsymbol{R}^{-1}\hat{\boldsymbol{R}}^{\frac{1}{2}} \\ \Leftrightarrow &\min_{\boldsymbol{Z},\boldsymbol{\Gamma},\boldsymbol{u},\{\boldsymbol{\sigma}\succeq 0\}} \mathrm{tr}(\boldsymbol{Z})+\mathrm{tr}\left(\hat{\boldsymbol{R}}^{-1}\boldsymbol{R}\right) \\ &\quad \text{subject to } \begin{bmatrix} \boldsymbol{Z} & \hat{\boldsymbol{R}}^{\frac{1}{2}} & \\ \hat{\boldsymbol{R}}^{\frac{1}{2}} & \boldsymbol{R} & \\ & & T(\boldsymbol{u}) \end{bmatrix} \geq 0 \quad . \end{aligned} \tag{11}$$

Substitute $\boldsymbol{R} = \boldsymbol{\Gamma}T(\boldsymbol{u})\boldsymbol{\Gamma}^{\mathrm{H}} + \mathrm{diag}(\boldsymbol{\sigma})$ into (11),

$$\begin{aligned} &\min_{\boldsymbol{Z},\boldsymbol{\Gamma},\boldsymbol{u},\{\boldsymbol{\sigma}\succeq 0\}} \mathrm{tr}(\boldsymbol{Z})+\mathrm{tr}(\boldsymbol{\Gamma}^{\mathrm{H}}\hat{\boldsymbol{R}}^{-1}\boldsymbol{\Gamma}T(\boldsymbol{u}))+\mathrm{Re}(\mathrm{diag}(\hat{\boldsymbol{R}}^{-1})^{\mathrm{H}})\boldsymbol{\sigma} \\ &\text{subject to } \begin{bmatrix} \boldsymbol{Z} & \hat{\boldsymbol{R}}^{\frac{1}{2}} & \\ \hat{\boldsymbol{R}}^{\frac{1}{2}} & \boldsymbol{\Gamma}T(\boldsymbol{u})\boldsymbol{\Gamma}^{\mathrm{H}}+\mathrm{diag}(\boldsymbol{\sigma}) & \\ & & T(\boldsymbol{u}) \end{bmatrix} \geq 0. \end{aligned} \tag{12}$$

We assume that $(\hat{\boldsymbol{u}},\hat{\boldsymbol{\sigma}})$ is the solution of the SDP problem (12) with initial/estimated $\hat{\boldsymbol{\Gamma}}$. The SDP is implemented using CVX with the SDPT3 solver [27-28]. Because $T(\hat{\boldsymbol{u}})$ is a positive semidefinite Toeplitz matrix, it can be decomposed to:

$$T(\hat{\boldsymbol{u}}) = \boldsymbol{V}\hat{\boldsymbol{P}}\boldsymbol{V}^{\mathrm{H}}, \tag{13}$$

where $\boldsymbol{V}$ denotes the eigenvector matrix and $\hat{\boldsymbol{P}}$ is a diagonal matrix and denotes the eigenvalues of $T(\hat{\boldsymbol{u}})$. The frequency parameter $\hat{\boldsymbol{f}}$ can be determined by the largest K eigenvalues. Then, the DOA estimates $\hat{\boldsymbol{\theta}}$ corresponding to $\hat{\boldsymbol{f}}$ as:

$$\hat{\boldsymbol{\theta}} = \arcsin(2\hat{\boldsymbol{f}}-1). \tag{14}$$

We know that $\mathrm{rank}(\boldsymbol{C}) = K \leq M-1$, and it is natural to determine that $\mathrm{rank}(T(\hat{\boldsymbol{u}})) = K \leq M-1$, which means that $T(\hat{\boldsymbol{u}})$ is rank-deficient. Therefore, the prior knowledge $K \leq M-1$ makes the solutions $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{P}}$ unique.

B. Gain error estimation

From (2), (3) and (4), it is clear that,

$$R_{m,l} = \begin{cases} \gamma_m\gamma_l^{*}C_{m,l}, & m \neq l \\ \gamma_m\gamma_m^{*}\sum_{k=1}^{K} p_k^2 + \sigma^2, & m = l \end{cases}, \tag{15}$$

when m=l, the main diagonal element $R(m,m)$ is that,

$$R(m,m) = \alpha_m^2\sum_{k=1}^{K} p_k^2 + \sigma^2. \tag{16}$$

$\hat{\boldsymbol{\sigma}}$ is estimated by (12), and the gain errors can be estimated as:

$$\hat{\alpha}_m = \mathrm{sqrt}\left(\frac{\hat{R}(m,m)-\hat{\sigma}^2}{\hat{R}(1,1)-\hat{\sigma}^2}\right), \quad m = 1,\ldots,M, \tag{17}$$

where $\hat{R}(1,1)$ is the first element of the main diagonal of $\boldsymbol{R}$ and is taken as a reference because $\alpha_1 = 1$.

C. Phase error estimation

In this section, we introduce a forward and backward space smoothing method to reduce the correlation between signal sources [24]. $\boldsymbol{J}_M$ is an M order exchange matrix where the reverse-diagonal elements are 1, and the rest of the elements are 0. We construct a new covariance matrix:

$$\tilde{\boldsymbol{R}} = \boldsymbol{R} + \boldsymbol{J}_M \boldsymbol{R}^* \boldsymbol{J}_M, \tag{18}$$

where $\boldsymbol{R}^*$ is the complex conjugate matrix of $\boldsymbol{R}$. The noise subspace $\boldsymbol{U} = [\boldsymbol{U}_{K+1}, \boldsymbol{U}_{K+2}, \ldots, \boldsymbol{U}_{K+M}]$ can be obtained by the characteristic decomposition of $\tilde{\boldsymbol{R}}$. It is worth noting that the space smoothing algorithm is a dimensionality reduction algorithm, which is essentially a process of restoring the rank of the covariance matrix. Therefore, in general, the dimension of the modified covariance matrix is smaller than that of the original covariance matrix, that is to say, the performance of estimating coherent signals is obtained by reducing the degree of freedom of the covariance matrix. However, in order to satisfy the dimension for estimating phase errors, we do not divide $\boldsymbol{R}$ and $\boldsymbol{R}^*$ into multiple sub-arrays to make $\tilde{\boldsymbol{R}}$ reach the optimal case of full rank. Thus, we may further improve the performance of the proposed method in the case of coherent signals in the future.

The phase errors can be estimated as [20]:

$$\hat{\boldsymbol{\phi}} = \text{angle}(\boldsymbol{w}), \tag{19}$$

$$\boldsymbol{w} = \frac{(\boldsymbol{B})^{-1}\boldsymbol{h}}{\boldsymbol{h}^{\text{T}}(\boldsymbol{B})^{-1}\boldsymbol{h}} \quad \boldsymbol{h} = [1, 0, \ldots, 0]^{\text{T}}, \tag{20}$$

$$\boldsymbol{B} = \sum_{k=1}^{K} [\text{diag}(\boldsymbol{a}(\hat{\boldsymbol{f}}_k))]^{\text{H}} \boldsymbol{U}\boldsymbol{U}^{\text{H}} \text{diag}(\boldsymbol{a}(\hat{\boldsymbol{f}}_k)), \tag{21}$$

where $\boldsymbol{\phi} \triangleq [\phi_1, \ldots, \phi_M]^{\text{T}}$ and angle($\bullet$) denotes the phase of a complex number.

Consequently, the proposed method is summarized as follows:

Step 1: Set $\boldsymbol{\Gamma}_0 = \boldsymbol{I}$ and $(\hat{\boldsymbol{u}}, \hat{\boldsymbol{\sigma}})$ is estimated by (12).

Step 2: $\hat{\boldsymbol{f}}$ and $\hat{\boldsymbol{\theta}}$ are estimated by (13) and (14), i.e., GLSM1.

Step 3: Gain errors are estimated by (17).

Step 4: Phase errors are estimated by (19).

Step 5: Set $\hat{\boldsymbol{\Gamma}} = \text{diag}([\hat{\alpha}_1 \text{e}^{\text{j}\hat{\phi}_1}, \ldots, \hat{\alpha}_M \text{e}^{\text{j}\hat{\phi}_M}]^{\text{T}})$, and $\hat{\boldsymbol{f}}$ and $\hat{\boldsymbol{\theta}}$ are re-estimated by (12), (13) and (14), i.e., GLSM2.

IV. CONNNECTION TO PRIOR ART

A. Relation with the ONGM method

The GLSM method proposed in this paper has some similarities with the ONGM in [24]. Both methods adopt the same covariance fitting criteria for DOA, a similar method for gain error estimation, and the same method for phase error estimation. However, the following three main differences make their performances very different. The first is that GLSM is based on the continuous frequency domain, while ONGM depends on the grid of the angular space. Note that the covariance matrix $\boldsymbol{R}$ is approximated in ONGM by discretizing the range of the angle, where it is no guarantee that the true DOAs are located on the grids. Thus, the modelling error, which is dependent on the grid density, is one possible reason for the inaccuracy of ONGM estimation. The second is that the DOA estimate of ONGM is obtained directly from the solution of the covariance fitting optimization problem [11], while GLSM obtains the frequency estimation by the Vandermonde decomposition of $T(\hat{\boldsymbol{u}})$. Therefore, the DOA estimation of ONGM is bound by the grid, which is referred to as an on-grid issue. The third is that parameters of ONGM are estimated via iterative processing, while parameters of GLSM are estimated with no iteration. Alternating iteration processing causes ONGM to easily fall into local optimality in the case of large phase errors.

B. Complexity analysis

In this section, we compare the computational complexity of the GLSM method with the ONGM method [24]. ONGM is an iterative algorithm whose computational complexity is equal to the complexity of each iteration times the number of iterations; therefore, the major computational complexity of the ONGM algorithm is $O\{(\tilde{K}^2M^3 + 2\tilde{K}M^4 + M^5)T\}$, where $\tilde{K}$ denotes the grid size and T denotes the number of iterations, which is difficult to quantify and varies in different scenarios. We define J_1 as the variable size and $J_2 \times J_2$ as the dimension of the positive semidefinite matrix in the semidefinite constraint of an SDP [11]. According to reference [29], we known that the SDP can be solved in $O(J_1^2 J_2^{2.5})$ flops. In the GLSM method, J_1 equals M^2 and J_2 equals M, so the complexity of the GLSM method is $O(M^{6.5})$. Obviously, the order on M for GLSM is higher than that for ONGM. However, the grid size $\tilde{K}$ is much greater than the number of sensors, M. Additionally, the computational complexity from phase errors is much higher in the ONGM method, which needs to execute $O(\tilde{K}TM^3)$ flops. In contrast, we need to run once to obtain the phase errors in the GLSM method. Therefore, the GLSM method can be faster than the ONGM method.

V. NUMERICAL SIMULATIONS

The gain errors $\{\alpha_m\}_{m=1}^{M}$ and phase errors $\{\phi_m\}_{m=1}^{M}$ of the sensors are generated by [21]:

$$\alpha_m = 1 + \sqrt{12}\sigma_\alpha \zeta_m, \tag{22}$$

$$\phi_m = \sqrt{12}\sigma_\phi \eta_m, \tag{23}$$

Where ζ_m and η_m are independent and identically distributed random variables distributed uniformly over [-0.5, 0.5]. Additionally, σ_α and σ_ϕ are the standard deviations of α_m and σ_ϕ, respectively [22]. In the following simulations, $\sigma_\alpha = 0.1$. For all Monte Carlo experiments, the number of trials is 200. It is known from (17) that the estimation of the gain errors is independent of the phase error and DOA, and the gain errors do not affect the accuracy of the DOA, so we omit the performance of the gain error estimation in the following simulations.

The RMSE of DOA and phase errors are measured in (24) and (25), respectively:

$$\text{RMSE}_\theta = \left[\frac{1}{200K}\sum_{k=1}^{K}\sum_{i=1}^{200}(\hat{\theta}_k^i - \theta_k)^2\right]^{1/2}, \tag{24}$$

$$\text{RMSE}_\phi = \left[\frac{1}{200M}\sum_{m=1}^{M}\sum_{i=1}^{200}(\hat{\phi}_m^i - \phi_m)^2\right]^{1/2}. \tag{25}$$

A. Effect of DOA separation

We consider K=2 uncorrelated sources. The sample number is 200, the signal-to-noise ratio (SNR) is 20 dB and $\sigma_\phi = 20^\circ$. The grid size of the ONGM method is 180. Figure 1 shows the space spectra of the WF method, the L1_SVD method, the ONGM method, and the GLSM method under two different DOA separation cases. The proposed method without calibration is referred to as GLSM1, and GLSM2 represents the proposed method with calibration. L1_SVD1 is without gain-phase errors, and L1_SVD2 is with gain-phase errors.

It can be seen from Fig. 1 (a) that when two signals are not close to each other ($\text{DOA} = [5^\circ\ 20^\circ]$), the performance of the L1_SVD method with gain-phase errors severely deteriorates and has only one sharp peak. In contrast, the other three methods can accurately estimate DOA. When the DOA separation of the two signals is reduced ($\text{DOA} = [17^\circ\ 20^\circ]$), Fig. 1 (b) shows that the WF method almost completely fails, and the ONGM method only has one peak. In contrast, the proposed method can still distinguish the two closed signals. In addition, the calibration technique of the proposed method improves the estimation accuracy (GLSM1 vs. GLSM2).

B. Effect of phase errors

We consider K=3 sources from directions $\boldsymbol{\theta} = [10, 25, 60]^{\mathrm{T}}$. The sample number is 200 and SNR is 20 dB. The grid size of the ONGM method is 180. Based on 200 Monte Carlo runs, the RMSE of DOA and gain-phase error estimates versus the standard deviation of the phase error σ_ϕ are obtained by the WF method, the ONGM method and the proposed method.

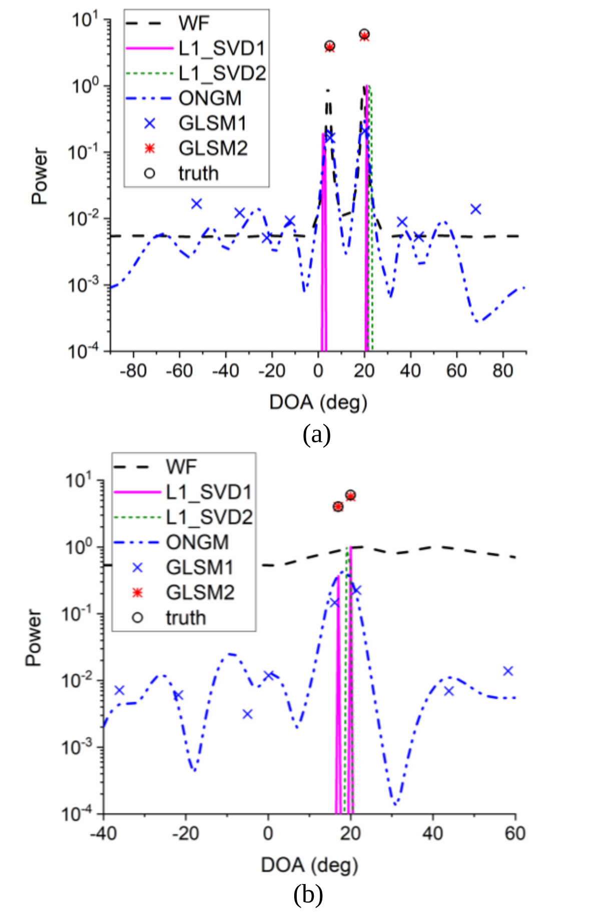
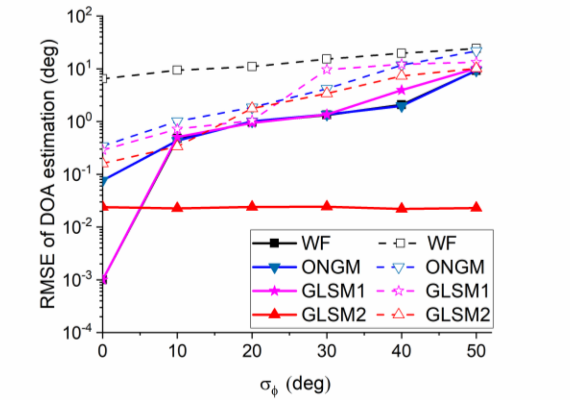
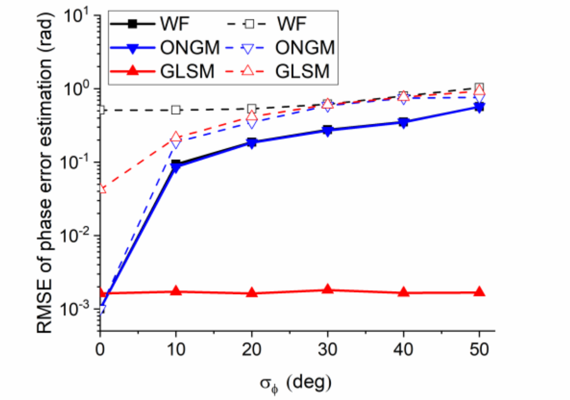


Fig. 1. Space spectrum of four algorithms: (a) $\text{DOA} = [5^\circ\ 20^\circ]$, and (b) $\text{DOA} = [17^\circ\ 20^\circ]$.

Figure 2 (a) and Fig. 2 (b) show that when the sources are uncorrelated, the ONGM method and the WF method are similar in the performance of DOA estimation and phase error estimation. Moreover, the performance of the two algorithms gradually deteriorates with increasing phase error because both algorithms easily fall into local optima when the phase errors are large. In contrast, the GLSM algorithm shows good stability to DOA and phase error estimation as the phase error increases from small to large, and the estimation accuracy is obviously higher than the first two algorithms. When the signal sources are coherent (source 2 is a replica of source 1), the WF algorithm is completely invalid in DOA estimation and phase error estimation. The estimation performance of the ONGM method and GLSM method are similar and obviously better than those of the WF method because GLSM and ONGM are based on the covariance fitting criterion, which is robust to the correlated sources, and both of them reduce the influence of signal correlation through the exchange matrix in the process of solving the phase error.



(a)



(b)

Fig. 2. RMSE versus σ_ϕ (a) RMSE of DOA estimation versus σ_ϕ, and (b) RMSE of phase error estimation versus σ_ϕ. (The dashed and solid plots represent the cases of coherent signal sources and uncorrelated signal sources, respectively).

C. Effect of SNR

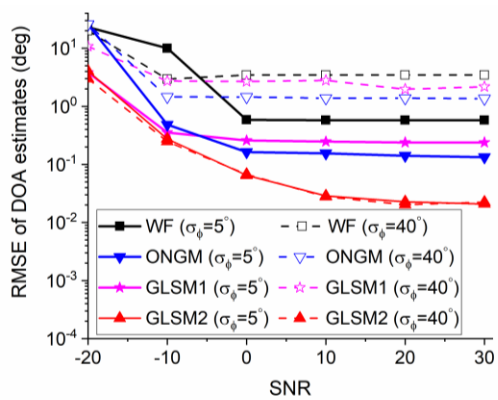
We consider three sources from directions by direction 10°, 25° and 60°. The number of samples is 200 and the varying SNR is from -20 dB to 30 dB. The grid number of the ONGM method is 180. Based on 200 Monte Carlo runs, the RMSE of DOA and phase error estimates versus SNR are obtained by the WF method, the ONGM method and the proposed method.

First, some simulations are performed to consider the effect of SNR in the case of uncorrelated signal sources. We show the results in two different phase error cases: $\sigma_\phi = 5^\circ$ and $\sigma_\phi = 40^\circ$. Figure 3 (a) and Fig. 3 (b) show that in the case of small phase error ($\sigma_\phi = 5^\circ$), the estimation performance of the three algorithms improves as the SNR increases. Moreover, the curve of the proposed method is constantly lower than that of the WF method and the ONGM method. When the phase error is large ($\sigma_\phi = 40^\circ$), the performance of the WF algorithm and ONGM algorithm is difficult to further improve with increasing SNR, and the GLSM algorithm still maintains good estimation performance.

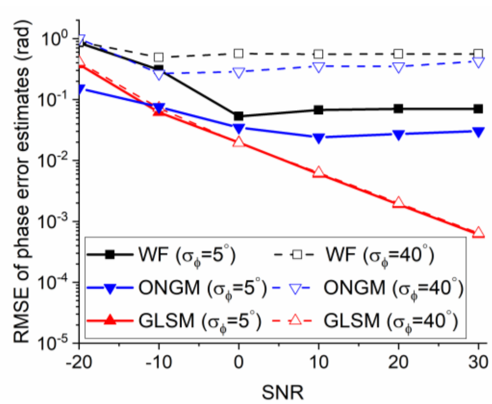
Then, some simulations are performed to consider the effect of SNR in the case of coherent sources (source 2 is a replica of source 1). The different simulation conditions from the uncorrelated sources case are two phase error cases ($\sigma_\phi = 5^\circ$ and $\sigma_\phi = 15^\circ$). Figure 4 (a) and Fig. 4 (b) show that in the case of coherent signal sources, the WF method is invalid regardless of how the SNR increases. In contrast, as the SNR increases, the phase error is smaller, the performance of the ONGM method and the GLSM method is better, and the GLSM method performs better than the ONGM method.

D. Effect of array length

We study the performance with respect to the array length M. Two uncorrelated sources impinge on the array with DOAs 10° and 25°. We set SNR=20 dB, $\sigma_\phi = 10^\circ$, and vary M from 5 to 40. Moreover, we consider two grid size cases (180 grids and 360 grids) for the ONGM method.



(a)



(b)

Fig. 3. RMSE versus SNR (the signal sources are uncorrelated). (a) RMSE of DOA estimates versus SNR. (b) RMSE of phase estimates versus SNR.

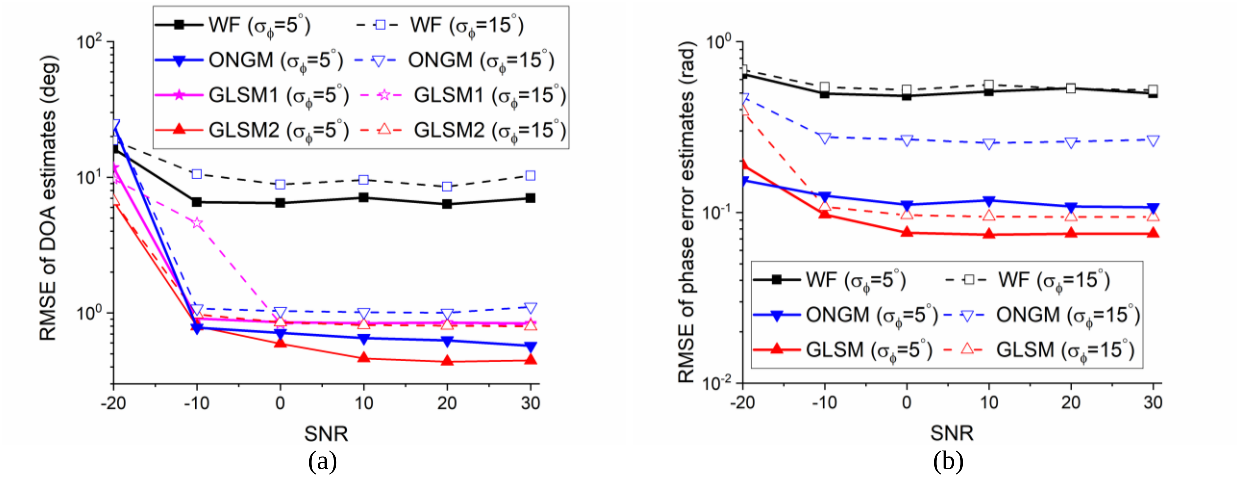


Fig. 4. RMSE versus SNR (Source 2 is a replica of source 1). (a) RMSE of DOA estimates versus SNR. (b) RMSE of phase error estimates versus SNR.

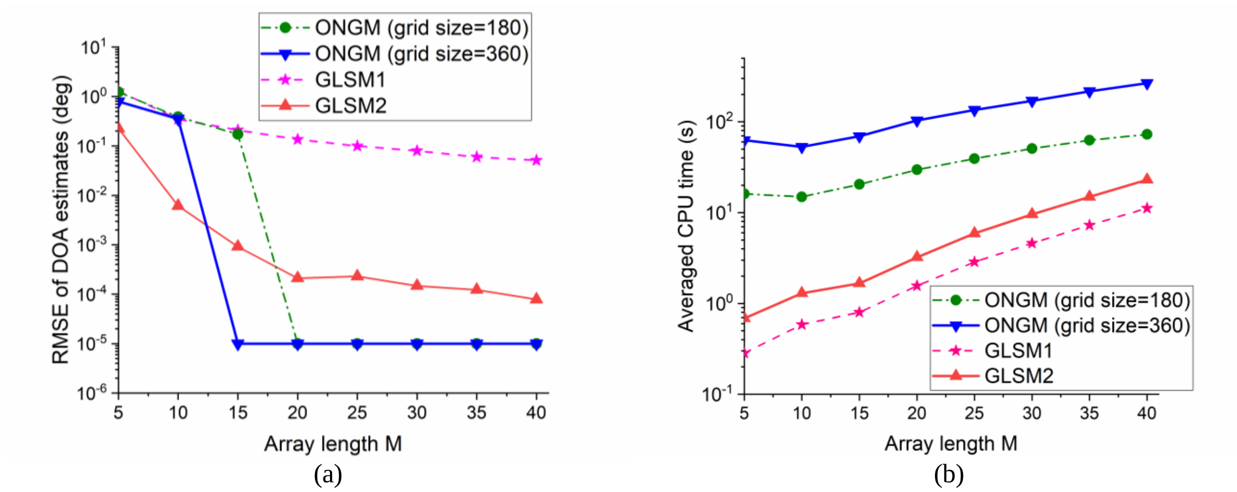


Fig. 5. (a) RMSE of DOA estimates of GLSM with an *M*-element ULA compared with ONGM. (b) CPU time usage of GLSM and ONGM.

Figure 5 (a) shows that the RMSEs of the GLSM method and the ONGM method improve with the array length. More specifically, when $M > 12$, the performance of the ONGM method with 360 grids is better than the GLSM method. Moreover, the more grids, the better the performance of the ONGM method is. Figure 5 (b) shows that the average CPU time usage of these two methods increases with an increasing number of array elements. However, the average CPU running time used by the GLSM method to complete one estimation is significantly lower than that of the ONGM method.

VI. CONCLUSION

In this paper, we propose a fast method to simultaneously estimate DOA and gain-phase errors in the continuous range without a grid mismatch problem. This method does not require the existence of a calibration source and calibration information. Monte Carlo runs show that utilizing the covariance fitting criterion and the SDP, the proposed algorithm overcomes the shortcomings of the ONGM method in the case of severe gain-phase error perturbations when signal sources are uncorrelated. Moreover, the proposed method has the advantages of low computational complexity and high resolution compared to existing methods.

REFERENCES

[1] R. Schmidt, “Multiple emitter location and signal parameter estimation,” *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276-280, 1986.

[2] R. Roy and T. Kailath, “ESPRIT-estimation of

signal parameters via rotational invariance techniques," *Adapt. Antennas Wirel. Commun.*, vol. 37, no. 7, pp. 224-235, 2009.

[3] S. A. Vorobyov, A. B. Gershman, and K. M. Wong, "Maximum likelihood direction-of-arrival estimation in unknown noise fields using sparse sensor arrays," *IEEE Trans. Signal Process.*, vol. 53, no. 1, pp. 34-43, 2005.

[4] J. Capon, "High-resolution frequency-wave-number spectrum analysis," *Proc. IEEE*, vol. 57, no. 8, pp. 1408-1418, 1969.

[5] D. Malioutov, M. Cetin, and A. S. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3010-3022, 2005.

[6] P. Stoica, P. Babu, and J. Li, "SPICE: A sparse covariance-based estimation method for array processing," *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 629-638, 2011.

[7] M. Muzi, N. Tedeschi, L. Scorrano, V. Ferrara, and F. Frezza, "Single-snapshot time-domain direction of arrival estimation under bayesian group-sparse hypothesis and vector sensor antennas," *Appl. Comput. Electromagn. Soc. J.*, vol. 33, no. 8, pp. 822-827, 2018.

[8] H. Cui, H. Duan, and H. Liu, "Off-grid DOA estimation using temporal block sparse Bayesian inference," in *2016 IEEE International Conference on Digital Signal Processing (DSP)*, 2016, pp. 204-207.

[9] 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, 2013.

[10] X. Zhang, T. Jiang, Y. Li, and X. Liu, "An off-grid DOA estimation method using proximal splitting and successive nonconvex sparsity approximation," *IEEE Access*, vol. 7, pp. 66764-66773, 2019.

[11] Z. Yang, L. Xie, and C. Zhang, "A discretization-free sparse and parametric approach for linear array signal processing," *IEEE Trans. Signal Process.*, vol. 62, no. 19, pp. 4959-4973, 2014.

[12] Z. Yang and L. Xie, "On gridless sparse methods for multi-snapshot direction of arrival estimation," *Circuits, Syst. Signal Process.*, vol. 36, no. 8, pp. 3370-3384, 2017.

[13] A. L. Swindlehurst and T. Kailath, "A performance analysis of subspce-based method in the presence of model errors, Part I: The MUSIC algorithm," *Signal Process. IEEE Trans.*, vol. 40, no. 7, pp. 1758-1774, 1992.

[14] A. L. Swindlehurst and T. Kailath, "A performance analysis of subspace-based methods in the presence of model errors, Part I: The MUSIC algorithm," *IEEE Trans. Signal Process.*, vol. 40, no. 7, pp. 1758-1774, 1992.

[15] A. Azarbar, G. R. Dadashzadeh, and H. R. Bakhshi, "2-D DOA estimation with matrix pencil method in the presence of mutual coupling," *Appl. Comput. Electromagn. Soc. J.*, vol. 27, no. 9, pp. 742-748, 2012.

[16] Q. Cheng, Y. Hua, and P. Stoica, "Asymptotic performance of optimal gain-and-phase estimators of sensor arrays," *IEEE Trans. Signal Process.*, vol. 48, no. 12, pp. 3587-3590, 2000.

[17] B. P. Ng, J. P. Lie, M. H. Er, and A. Feng, "A practical simple geometry and gain/phase calibration technique for antenna array processing," *IEEE Trans. Antennas Propag.*, vol. 57, no. 7, pp. 1963-1972, 2009.

[18] W. Peng, Y. Gao, Y. Qu, and C. Guo, "Array calibration with sensor gain and phase errors using invasive weed optimization algorithm," in *2017 Sixth Asia-Pacific Conference on Antennas and Propagation (APCAP)*, pp. 1-3, 2017.

[19] A. Paulraj, R. Roy, and T. Kailath, "A subspace rotation approach to signal parameter estimation," *Proc. IEEE*, vol. 74, no. 7, pp. 1044-1046, 1986.

[20] A. J. Weiss and B. Friedlander, "Eigenstructure methods for direction finding with sensor gain and phase uncertainties," *Circuits, Syst. Signal Process.*, vol. 9, no. 3, pp. 271-300, 1990.

[21] A. Liu, G. Liao, C. Zeng, Z. Yang, and Q. Xu, "An eigenstructure method for estimating DOA and sensor gain-phase errors," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 5944-5956, 2011.

[22] S. Cao, Z. Ye, D. Xu, and X. Xu, "A hadamard product based method for DOA estimation and gain-phase error calibration," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 49, no. 2, pp. 1224-1233, 2013.

[23] H. Liu, L. Zhao, Y. Li, X. Jing, and T. K. Truong, "A sparse-based approach for DOA estimation and array calibration in uniform linear array," *IEEE Sens. J.*, vol. 16, no. 15, pp. 6018-6027, 2016.

[24] W. Peng, C. Guo, M. Wang, and Y. Gao, "An improved gain-phase error self-calibration method for robust DOA estimation," *Int. J. Microw. Wirel. Technol.*, pp. 1-9.

[25] P. Stoica, P. Babu, and J. Li, "New method of sparse parameter estimation in separable models and its use for spectral analysis of irregularly sampled data," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 35-47, 2011.

[26] C. R. Rojas, D. Katselis, and H. Hjalmarsson, "A note on the SPICE method," *IEEE Trans. Signal Process.*, vol. 61, no. 18, pp. 4545-4551, 2013.

[27] K.-C. Toh, M. J. Todd, and R. H. Tütüncü, “SDPT3—A MATLAB software package for semidefinite programming, version 1.3,” *Optim. Methods Softw.*, vol. 11, no. 1-4, pp. 545-581, 1999.

[28] M. Grant, S. Boyd, and Y. Ye, “CVX: Matlab software for disciplined convex programming,” 2009.

[29] L. Vandenberghe and S. Boyd, “Semidefinite programming,” *SIAM Rev.*, vol. 38, no. 1, pp. 49-95, 1996.

Wencan Peng was born in Hubei Province, China, in 1987. She received the B.S. degree in School of Measuring and Optical Engineering, Nanchang Hangkong University in Nanchang city, China, in 2010. Then, she received the M.S. degree in School of Electronics and Information, Northwestern Polytechnical University in Xi’an city, China, in 2013. Now, She is pursuing the Ph.D. degree in School of Electronics and Information, Northwestern Polytechnical University. Her research interests include: array signal processing and array calibration.

Chenjiang Guo was born in Shannxi Province, China, in 1963. CIE Senior Member, Antenna Society Committee Member. He received the B.S., M.S. and Ph.D. in School of Electronics and Information, Northwestern Polytechnical University in Xi’an City, China, in 1984, 1987 and 2007 respectively. He is a Professor in School of Electronics and Information NWPU. He has published more than 140 research papers. He research interests includes: array signal processing, theory and design of antenna.

Min Wang (IEEE Member), received the B.S. degree from Xidian Univ. China in 2000; M.S. and Ph.D. degree in Signal and Information Processing from Xidian Univ., Xi’an China, in 2003 and 2005 respectively. He is working with National Lab of Radar Signal Processing in Xidian University. His research interests include sparse signal processing, miliwave/Terahertz radar and high resolution radar imaging.

Yuteng Gao was born in Shannxi Province, China, in 1988. He received the B.S. and M.S degree in School of Electronics and Information, Northwestern Polytechnical University in Xi’an City, China, in 2010 and 2013 respectively. He is presently working on his Ph.D. degree in School of Electronics and Information, Northwestern Polytechnical University in Xi’an City, China. His research interests include milimeter wave, antenna design and radar imaging.

Xia Zhao was born in Ningxia, China, in 1978. She is now an Associate Professor in North Minzu University, and pursuing her Ph.D in Electromagnetic Field and Microwave Technology at Northwestern Polytechnical University. Her research interests concern: array signal processing and target localization.