IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received December 16, 2020, accepted December 30, 2020, date of publication January 18, 2021, date of current version January 29, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3050602

Direction of Arrival Estimation by Matching Pursuit Algorithm With Subspace Information

YANG ZHAO[1], SI QIN[2], YI-RAN SHI[1], (Member, IEEE), AND YAO-WU SHI[1]
[1]College of Communication Engineering, Jilin University, Changchun 130025, China
[2]Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China

Corresponding author: Yi-Ran Shi (shiyiran@jlu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61571462, and in part by the Jilin Provincial Natural Science Foundation of China under Grant 20170101207JC.

ABSTRACT Traditional orthogonal matching pursuit (OMP) algorithms for direction of arrival (DOA) estimation suffer from poor angular resolution and noise suppression. In this paper, we analyze the reason why the OMP algorithm has difficulties in resolving closely separated DOAs and conclude that it lies in the rules of support detection. Moreover, we propose a solution to this problem via developing the connection between the sparse reconstruction class algorithm and the subspace algorithm from the structure of the redundant dictionary. Based on the framework of the matching pursuit (MP) algorithm, the effective information of the signal and noise subspaces is integrated, and a noise subspace reprojection orthogonal matching pursuit (NSRomp) algorithm for DOA estimation is proposed. By adopting signal subspaces to reconstruct the original signal, the proposed NSRomp can reduce both the influence of noise on the selection of the support set and the computing time. By implementing the minimum norm method to optimize the noise subspace into a vector, which corrects the selection rules of the support set during each iteration, the angular resolution of the proposed algorithm is improved. From the simulation results, when the signal to noise ratio (SNR) is lower than or near 0, the angular resolution can be improved by $> 15°$ using OMP algorithms to by $< 5°$ using the proposed NSRomp algorithm.

INDEX TERMS Bearing estimation, compressed sensing, DOA, MMVOMP, multiple measurement vector, MUSIC.

I. INTRODUCTION

The rapid development of wideband and ultrawideband signal processing technology demands simple and efficient methods to obtain information by sampling data. In many practical systems, information of interest is often sparse in the vast observable signal space and thus requires extensive until the recent applications of sparse signal representation (SSR) and compressed sensing (CS) theory. The theory notes that [1]–[3] if a signal is sparse or sparse after a certain transformation, the high-dimensional signal can be projected into a low-dimensional space and reconstructed from a small set of low-dimensional data with high probability. CS theory affects many research fields, including radar image processing [4]–[6], blind source separation [7]–[9], sensor networks [10], [11], and Internet of Things communication [12]. Because signal DOAs are substantially sparse in all spatial domains, DOA estimation was extensively studied using SSR. In 2005, Malioutov *et al*. [13] proposed an $l_1$svd algorithm that introduced SSR into DOA estimation. In the one-dimensional DOA estimation system, a sparse recovery model (SRM) was built, and a l_0-norm constrained optimization problem was solved by the convex optimization theory. The proposed $l_1$svd algorithm can be used with any structure array and correlation signal with satisfactory angular resolution, which improved the base pursuit (BP) algorithms with their variants, such as least absolute shrinkage and selection operator(LASSO) [14] or basis pursuit de-noising(BPDN) [15], adaptive LASSO [16] and weight LASSO [17]. These algorithms commonly referred to BP algorithm, which is similar to the l_1 norm and is used to approximate the l_0 norm such that optimization can be relaxed to l_1 norm-constrained convex optimization, and global optimality can be derived. However, these methods

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

suffer from biased estimation and complex calculations [18]. When the correlation between the different components of the observation model is strong, the l_1 norm cannot guarantee that the solution obtained is completely consistent with the real signal model. The $l_p(0 < p < 1)$ optimization results have better performance than BP because the l_p norm is a closer approximation to the l_0 norm [19], [20]. However, the complexity of optimizing the l_p norm greatly limits the implementation of the related methods. Compared to the $l_p(0 < p \leq 1)$ norm method, which has to solve a time-consuming optimization problem, the matching pursuit (MP) algorithm was proposed by Mallat and Zhang [21], [22] to reconstruct the sparsest target signal through linear measurements. Following Mallat and Zhang's research, OMP [23], MMVOMP [24], ROMP [25] and other improved algorithms were proposed. Iteration based MP class algorithms are greedy algorithms which cannot guarantee global optimality, but run quickly with fewer snapshots, and make MP class algorithms can estimate multidimensional parameters. Currently, MP based DOA estimation algorithms have primarily been proposed by Karabulut [26], Cotter [27] and Wang and Wu *et al.* [28]. However, existing MP based DOA estimation algorithms have several drawbacks to constrained their wide applications:

- Multiple snapshots are employed to for calculation but are not fully used to suppress noise;
- The angular resolution is relatively low, and distinguishing closely separated DOAs is difficult.

Fortunately, the classic eigensubspace methods, which are regarded as super-resolution algorithms, yield better spatial angle resolutions. The basic concept behind these methods is that SVD decomposition is employed to divide the data space into signal subspaces and noise subspaces, and then obtain the DOA estimation via the rotational invariance technique (e.g., ESPRIT [29]) or spectral peak searching technique (e.g., multiple signal classification (MUSIC [30])). Recently, [31]–[35] have been reported to combine sparse and subspace methods to compensate for their limitations. However, these combined methods were primarily concentrated on BP algorithms with complex second order cone programming (SOCP) with the requirement of heavy calculation; thus, these methods could not be applied to scenarios with real-time requirements. In this paper, motivated by the subspace algorithm, we describe a high-resolution NSRomp algorithm for DOA estimation that requires fewer calculation and exhibits complexity. We used SVD to divide the data space in the autocorrelation domain into the signal and noise subspaces. The obtained data from the signal subspace were considered to be the initial residual signal (i.e., the input) of the OMP algorithm, and with the noise subspace, the atom selection rules were modified at each iteration of the algorithm. The proposed NSRomp algorithm exploits the data of the two subspaces without data loss, makes it more efficient than merely using data from a single subspace. The contributions of this paper are summarized as follows:

- Through theoretical analysis, we conclude that the MP class algorithms are limited by the Rayleigh limit, which prevents them from resolving closely separated DOAs.
- The connection between the sparse reconstruction algorithm and the subspace algorithm from the structure of the redundant dictionary is built.
- The angular resolution of the MP class algorithm is improved by amending the rules of support detection.
- We develop a feasible algorithm based on the signal and noise subspaces via the OMP framework.

The remainder of the paper is organized as follows. In Section II, the sparse model of DOA estimation is shown. Then, based on the flow of the OMP algorithm, we discuss the reason of the degradation of the angular resolving resolution furthermore this problem is solved. After that, the minimum norm method is introduced, and NSRomp for DOA estimation is implemented. Section III presents the numerical simulation results, and Section IV draws conclusions.

Notations used in this paper are as follows. Vectors and matrices are presented by lowercase boldface and uppercase boldface, respectively. $\mathbf{I}_M$ is an $M \times M$ identity matrix. $(\cdot)^*$, $(\cdot)^T$ and $(\cdot)^H$ denote the conjugate, transpose, and conjugate transpose of the matrix, respectively. $vec(\cdot)$ denotes the vectorization operator that turns a matrix into a vector. $diag(a)$ denotes a diagonal matrix that uses the elements of a as its diagonal elements. $Tr(\cdot)$ represents the trace of the matrix. $|\cdot|$, $\|\cdot\|_2$ and $\|\cdot\|_F$ denote the cardinality, the l_2 norm and the Frobenius norm of the matrix, respectively. $\langle\cdot\rangle$ represents the inner product.

II. PRELIMINARIES AND METHODS

A. SPACE MODEL OF DOA ESTIMATION

In this subsection, the DOA estimation problem is stated. A uniform linear array (ULA) of N sensors can be illuminated by $K(K < N)$ narrowband far-field and uncorrelated sources whose DOAs θ_k satisfies $-\pi/2 \leqslant \theta_k \leqslant \pi/2$ for $k = 1, 2, \ldots, K$. Then, the received data vector at the t-th snapshot can be expressed by:

$$\mathbf{X}(t) = \mathbf{A}(\theta)\mathbf{S}(t) + \mathbf{N}(t), \quad t = 1, 2, \ldots, L. \tag{1}$$

where L is the number of snapshots, $\mathbf{n}(t)$ is the additive Gaussian white noise with zero mean, the covariance matrix $\sigma_n^2\mathbf{I}$ is uncorrelated to the source signals, $\mathbf{A}(\theta) = [\mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_K)]$ is the manifold matrix, and $\mathbf{a}(\theta_k)$ is the steering vector of the k-th source, which can be expressed by:

$$\mathbf{a}(\theta_k) = \left[1, e^{j2\pi d \sin(\theta_k)/\lambda}, \ldots, e^{j2\pi d(N-1)\sin(\theta_k)/\lambda}\right]^T. \tag{2}$$

where λ is the wavelength; d is the distance between two adjacent sensors ($d \leq \lambda/2$)); and K is a given number of sources.

Remark 1: The source number K is typically unknown in practice, and an estimate of K can be obtained by eigenvalue decomposition, information theory, Gai's circle method, regular correlation method, and other methods.

Based on [13], under the receiving range of the ULA ($-\pi/2 \leq \theta_k \leq \pi/2$), all possible spatial angles are divided

into N_s parts on average:

$$\Omega = \left\{ \bar{\theta}_i \mid -\pi/2 \le \bar{\theta}_i \le \pi/2, \bar{\theta}_i < \bar{\theta}_{i+1}, \quad i = 1, 2, \ldots, N_s \right\}. \tag{3}$$

where N_s is much greater than K. Then, the virtual steering vector matrix as a dictionary matrix is constructed from Ω:

$$\Phi(\bar{\theta}) = \left[\mathbf{a}\left(\bar{\theta}_1\right), \mathbf{a}\left(\bar{\theta}_2\right), \ldots, \mathbf{a}\left(\bar{\theta}_{N_s}\right) \right]$$

where $\mathbf{a}\left(\bar{\theta}_i\right)$ and $\mathbf{a}\left(\theta_k\right)$ can be obtained from (2). If $\theta_k \in \Omega, k = 1, 2, \ldots, K$, we can derive the sparse recovery model from (1) as

$$\mathbf{X}(t) = \Phi(\bar{\theta})\bar{\mathbf{S}}(t) + \mathbf{N}(t), \quad t = 1, 2, \ldots, L. \tag{4}$$

where: $\bar{\mathbf{S}}(t) = \left[\bar{\mathbf{s}}_1(t), \bar{\mathbf{s}}_2(t), \ldots, \bar{\mathbf{s}}_{N_s}(t) \right]^T$,

$$\bar{\mathbf{s}}_i(t) = \begin{cases} \bar{\mathbf{s}}_k(t), & \text{when } \bar{\theta}_i = \theta_k \\ 0, & \text{when } \bar{\theta}_i \ne \theta_k \end{cases} \quad i = 1, 2, \ldots, N_s, k = 1, 2, \ldots, K \Bigg\}. \tag{5}$$

The objective of this work is to find a signal $\bar{\mathbf{S}}(t)$ with a sparse signal recovery algorithm, and the DOA estimation is accomplished based on the position of nonzero elements.

B. INSUFFICIENT ANGLE RESOLUTION OF MP CLASS ALGORITHMS

In this subsection, we discuss why MP class algorithms can be sensitive to closely spaced sources from the flow of the OMP algorithm. It can be explained as follows: First, the OMP algorithm regards the output signal of the sparse array model as the initial value of the residual error signal and then correlates the current residual signal with the column vectors in the redundant dictionary matrix. Second, a basic atom with the largest correlation is selected as an alternative basis. Third, the components of the selected base atoms are removed from the current residual error signal to obtain a new residual error signal. The process is then repeated until termination conditions are satisfied.

The support set based on the correlation maximum principle is chosen by the general MP class algorithm as:

$$\lambda_k = \arg \max_{i=1,2\ldots,N_s} \left\{ \left| \left\langle \mathbf{a}\left(\bar{\theta}_i\right), \mathbf{r}_{k-1} \right\rangle \right| \right\}. \tag{6}$$

where $\mathbf{r}_{k-1}$ is the residual of the $(k-1)$-th iteration, and λ_k is the index of the selected atom in the dictionary at the k-th iteration.

Assuming that the array receives a plane wave signal at the angle of θ_0, the wave shape of the signal is $\mathbf{S}(t)$, and θ_0 lies on the grid Ω (only if the grid density of the dictionary is sufficiently large). Then $\mathbf{a}(\theta_0) = \mathbf{a}\left(\bar{\theta}_0\right)$, where $\mathbf{a}\left(\bar{\theta}_0\right)$ is numbered as λ_0 in $\Phi(\bar{\theta})$. Ideally, the λ_0-th element of the sparse vector $\bar{\mathbf{S}}(t)$ is $\mathbf{S}(t)$, and the remaining elements are all equal to 0. From (4), the initial residual error is:

$$\mathbf{r}_0 = \mathbf{X}(t) = \Phi(\bar{\theta})\bar{\mathbf{S}}(t) + \mathbf{N}(t) = \mathbf{a}\left(\bar{\theta}_0\right)\mathbf{s}(t) + \mathbf{n}(t). \tag{7}$$

Based on (6), the inner products of $\mathbf{a}\left(\tilde{\theta}_j\right)$ and $\mathbf{r}_0$ are:

$$\left\langle \mathbf{a}\left(\bar{\theta}_i\right), \mathbf{r}_0 \right\rangle = \mathbf{s}^H(t)\mathbf{a}^H\left(\bar{\theta}_0\right)\mathbf{a}\left(\bar{\theta}_i\right) + \mathbf{n}^H(t)\mathbf{a}\left(\bar{\theta}_i\right). \tag{8}$$

where $i = 1, 2, \ldots, N_s$, and $\mathbf{n}^H(t)\mathbf{a}\left(\bar{\theta}_i\right)$ can be neglected because the noise $n(t)$ is nearly uncorrelated to $\mathbf{a}\left(\bar{\theta}_i\right)$. When the λ_0-th angle is explored, $\left\langle \mathbf{a}\left(\bar{\theta}_i\right), \mathbf{r}_0 \right\rangle$ reaches its maximum $\mathbf{s}(t)N + \mathbf{a}\left(\bar{\theta}_0\right)N_L$ due to the coherent addition of the signal from the direction of θ_0. The optimal λ_0 is then identified.

If we assume that the array receives two plane wave signals at angles θ_0 and θ_1, respectively, then the wave shapes of signals are $\mathbf{s}_0(t), \mathbf{s}_1(t)$, while θ_0 and θ_1 lie on grid Ω only if the grid density of the dictionary is sufficiently large. Then, $\mathbf{a}(\theta_i) = \mathbf{a}\left(\bar{\theta}_i\right)$, where $\mathbf{a}\left(\bar{\theta}_i\right)$ is numbered as $\lambda_i \quad (i = 0, 1)$. Because the signal wave shapes do not contain information about direction, we assume that the wave shapes of the two signals are sufficiently similar that $\mathbf{s}_0(t) \approx \mathbf{s}_1(t) = \mathbf{s}(t)$. Based on (6), the inner products of atoms and $\mathbf{r}_0$ are given by:

$$\left\langle \mathbf{a}\left(\bar{\theta}_i\right), \mathbf{r}_0 \right\rangle = \left[\mathbf{a}^H\left(\bar{\theta}_i\right)\mathbf{a}\left(\bar{\theta}_0\right) + \mathbf{a}^H\left(\bar{\theta}_i\right)\mathbf{a}\left(\bar{\theta}_1\right) \right] \mathbf{s}(t) + \mathbf{a}^H\left(\bar{\theta}_i\right)\mathbf{n}(t). \tag{9}$$

If all dictionary atoms are normalized as $\left\|\mathbf{a}\left(\bar{\theta}_i\right)\right\|_2 = 1$. The term $\mathbf{a}^H\left(\bar{\theta}_i\right)\mathbf{a}\left(\bar{\theta}_0\right) + \mathbf{a}^H\left(\bar{\theta}_i\right)\mathbf{a}\left(\bar{\theta}_1\right)$ in (9) will vary with $\bar{\theta}_i$ and form a wave shape $\mathbf{\Lambda}\left(\bar{\theta}_i\right)$ composed of discretized $\bar{\theta}_i = 1, 2, \ldots, N_s$.

Intuitively, $\mathbf{\Lambda}\left(\bar{\theta}_i\right)$ is generated by the superposition of the two sinusoidal peaks at $\bar{\theta}_0, \bar{\theta}_1$. When θ_0 and θ_1 are sufficiently far in the angular space, $\left\langle \mathbf{a}\left(\bar{\theta}_i\right), \mathbf{r}_0 \right\rangle$ reaches its maximum successfully with the λ_0 th or λ_1 th atoms, and the OMP algorithm functions correctly. However, when θ_0 and θ_1 are close in the angular space, $\mathbf{\Lambda}\left(\bar{\theta}_0\right)$ and $\mathbf{\Lambda}\left(\bar{\theta}_1\right)$ will interfere and form a new wave shape. Then, affected by the term $\mathbf{a}^H\left(\bar{\theta}_i\right)\mathbf{n}(t)$, $\left\langle \mathbf{a}\left(\bar{\theta}_i\right), \mathbf{r}_0 \right\rangle$ will reach multiple optimal solutions between θ_0 and θ_1, and the OMP algorithm fails.

If we investigate the mutual coherence of $\Phi(\bar{\theta})$, it is easy to predict this result. Let the pairwise coherence between the k-th and l-th columns be [26]:

$$\mu(k, l) = \left| \frac{\left\langle \mathbf{a}\left(\bar{\theta}_k\right), \mathbf{a}\left(\bar{\theta}_l\right) \right\rangle}{\left|\mathbf{a}\left(\bar{\theta}_k\right)\right| \left|\mathbf{a}(\bar{\theta})\right|} \right|. \tag{10}$$

Definition 1: The mutual coherence of $\Phi(\bar{\theta})$ is the maximum pairwise coherence of columns:

$$\mu(\Phi(\bar{\theta})) = \max_{k \ne l} \mu(k, l). \tag{11}$$

If let the number of sensor elements be $M = 10$ and the grid spacing be 1°, $\mu(\Phi(\tilde{\theta})) = 0.999999758842$. If the difference between θ_0 and θ_1 is less than 1°, we have to account for at least 7 decimal places to guarantee the validity of the OMP algorithm.

For quantitative analysis, we assume that θ_0 is already given. Then, the module of $\mathbf{a}^H\left(\bar{\theta}_i\right)\mathbf{a}\left(\bar{\theta}_0\right)$ can be considered

to be a function of $\bar{\theta}_i$

$$\mathbf{F}\left(\bar{\theta}_i\right) = \left|\left\langle \mathbf{a}^H\left(\bar{\theta}_i\right), \mathbf{a}\left(\bar{\theta}_0\right)\right\rangle\right| = \left|\sum_{i=1}^{N-1} e^{-j\frac{2\pi d}{\lambda} i \sin\bar{\theta}_i} e^{j\frac{2\pi d}{\lambda} i \sin\bar{\theta}_0}\right| = N \sin c\left(\boldsymbol{\varphi}_j - \boldsymbol{\varphi}_0\right). \quad (12)$$

where $\boldsymbol{\varphi}_i = \frac{2\pi d}{\lambda}\sin\bar{\theta}_i$, $\sin c(\boldsymbol{\varphi}) = \frac{\sin\frac{N_{\boldsymbol{\varphi}}}{2}}{\sin\frac{\boldsymbol{\varphi}}{2}}$. From the first zero position of the $\sin c$ function in (12), the beam width can be given by the half power point (3 dB) of $\mathbf{F}\left(\bar{\theta}_i\right)$, $\frac{\lambda}{Nd}$. Additionally, if we assume θ_1 is already given, then $\left|\mathbf{a}^H\left(\bar{\theta}_i\right)\mathbf{a}\left(\bar{\theta}_1\right)\right|$ is a function of $\bar{\theta}_i$. Drawing the curve can yield the beam width in the direction of θ_1. The beam widths can be considered to be the valid range. When $\left|\bar{\theta}_0 - \bar{\theta}_1\right| \leq \frac{\lambda}{Nd}(rad)$, the valid ranges of signals θ_0 and θ_1 will interfere with each other and lead to the incompetence of the OMP algorithm. Thus, we can conclude that the angular resolution of the OMP algorithm based on (6) is $\frac{\lambda}{Nd}(rad)$, which is the so-called Rayleigh limit and proportional to the array aperture Nd.

When there are two or more spatial distances approaching DOAs within a beam width, the method of support set selection based on the principle of maximum correlation fails. Due to the strong correlation between the adjacent atoms in the redundant dictionary, when there are two or more closely separated DOAs in the system, one to one mapping between the maximum point of (9) and the true value of DOA is broken. Thus, only the approximate range of DOA estimation could be given; however, accurate estimation was not allowed.

Remark 2: With regard to the iterative algorithm (e.g., matching pursuit), the later iteration results would be directly affected by the error of the supporting set selected at the previous iteration. When it is used for DOA estimation, this mismatch would lead to estimation bias and even missing estimation.

Thus, (6) is the fundamental reason why the OMP algorithm fails to break the Rayleigh limit. To improve the angular resolution of the algorithm, it is necessary to improve the selection rules of the support set.

C. BREAKING THROUGH THE RAYLEIGH LIMIT

The MP algorithm is designed to sequentially find support for an unknown signal in the dictionary, and then reconstruct the signal using the linear combination of the set. The dictionary in the sparse recovery model of DOA estimation in the literature typically spans the steering vectors throughout the space. The following lemmas can be obtained using this sparse model:

Lemma 1: For all sparse recovery algorithms that use the array manifold matrix to generate a redundant dictionary, the DOA problem is equivalent to find the correct steering vector of the system in the dictionary if the grid distribution is sufficiently detailed to make the true value of DOAs lie on the grid of the redundant dictionary.

Using Lemma 1, the relationship between the compressed sensing algorithms and the subspace algorithms was developed.

Remark 3: According to the classic subspace theory, the space of all steering vectors was the same as the signal subspace; this scenario was equivalent to reduce the search scope and to find the steering vectors in the signal subspace compared to perform that in the full angle space.

Based on the conclusion, the signal subspace was used as the input of the proposed NSRomp algorithm. Similar ideas were also used in the L1svd algorithm.

Remark 4: Unlike the subspace algorithm, L1svd decomposes array output sample data $\mathbf{X}$ in (1) by SVD. When $L \gg M$, using the eigenvalue decomposition of the sample covariance matrix $\mathbf{X}\mathbf{X}^H$ to replace it can effectively reduce computation time.

Inspired by the subspace algorithms, the orthogonality of two subspaces can be used to establish a new support set selection rule for the MP algorithm. Theorem 1 can be obtained from Lemma 1.

Theorem 1: If the DOA estimated in an iteration of the matching pursuit class algorithm is correct, the noise subspace and the atom selected at this iteration must be orthogonal.

Proof: Let the k -th iteration of the OMP algorithm obtain the support set as the atom $\mathbf{a}_k$ in the dictionary. If the index in the dictionary where $\mathbf{a}_k$ is located is the exact value of the DOA, $\mathbf{a}_k$ is the steering vector. Then, $\mathbf{a}_k$ must be in the space $\mathbf{A}$, which is composed of all steering vectors, $\mathbf{A}$ is the same space as the signal subspace $\mathbf{V}_S$, and $\mathbf{V}_S$ is orthogonal to the noise subspace $\mathbf{V}_N$; thus, $\mathbf{a}_k$ and $\mathbf{V}_N$ are orthogonal;

Conversely, if $\mathbf{a}_k$ and $\mathbf{V}_N$ are not orthogonal, $\mathbf{a}_k$ is not in $\mathbf{V}_S$. Therefore, $\mathbf{a}_k$ is not a steering vector, and the index of $\mathbf{a}_k$ in the dictionary is not the exact DOA value. Therefore, $\mathbf{a}_k$ is not a steering vector, and the index of $\mathbf{a}_k$ in the dictionary is not an accurate DOA value.

Based on [34], the i-th column of the data covariance matrix $\mathbf{R}_s$ can be remodeled as (4):

$$\mathbf{R}_s(:, i) = \mathbf{r}_i = E\left[\mathbf{X}(t)\mathbf{x}_i^*(t)\right] = \Phi(\bar{\theta})\mathbf{b}_i + \sigma_n^2\mathbf{e}_i. \quad (13)$$

where $i = 1, 2, \ldots, N$; $\mathbf{b}_i$ is a sparse coefficient vector of $\mathbf{r}_i$ under dictionary $\Phi(\bar{\theta})$; and $\mathbf{e}_i$ is a $N_s \times 1$ column vector with the i-th element being 1 and the remainder being 0. Changing (13) to matrix form based on all columns of $\mathbf{R}_s$ yields:

$$\mathbf{R}_s = \Phi(\bar{\theta})\mathbf{B} + \sigma_n^2\mathbf{I}_N. \quad (14)$$

where $\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_N]$ and $\mathbf{b}_i (i = 1, 2, \ldots, N)$ should have identical sparse structure (i.e., the nonzero elements of each ideal should appear in the same rows of $\mathbf{B}$). Because we are only interested in sparse distribution, (13) can be considered to be the same solution, and (14) can be considered an MMV problem composed of multiple snapshots. $\mathbf{R}_s$ in (14) is a positive definite Hermite matrix $n \times n$. The eigenvalue decomposition of $\mathbf{R}_s$ can be represented as:

$$\mathbf{R}_s = \mathbf{U}\Lambda\mathbf{V} = \mathbf{U}_S\Lambda_S\mathbf{U}_S^H + \mathbf{V}_N\Lambda_N\mathbf{V}_N^H. \quad (15)$$

where $\mathbf{U}_S \in \mathbb{C}^{N\times K}$ denotes the signal subspace collected from the eigenvectors of the K largest eigenvalues contained in the diagonal of $\mathbf{\Lambda}_S$. The remaining $N-K$ eigenvalues of $\mathbf{R}_s$ in the diagonal of $\mathbf{\Lambda}_N$ form the noise subspace $\mathbf{V}_N$.

Combining the relationship shown in (14) and (15) a new signal output equation can be obtained as:

$$\mathbf{X}_s = \Phi(\bar{\theta})\mathbf{B} + \mathbf{N}_L. \tag{16}$$

where $\mathbf{N}_L = \Delta\mathbf{R} - \mathbf{V}_N \Lambda_N \mathbf{V}_N^H$ is the noise error term; $\Delta\mathbf{R}$ is the error caused by approximating $\widehat{\mathbf{R}}_s = \sum_{t=1}^{L} \mathbf{x}(t)\mathbf{x}(t)^H$ with $\mathbf{R}_s$; and $\mathbf{X}_s = \mathbf{U}_S\mathbf{\Lambda}_S\mathbf{U}_S^H - \boldsymbol{\sigma}_n^2\mathbf{I}_N$ is the signal subspace component.

Remark 5: Intuitively, using (16) to instead of (1) can reduce the number of snapshots from L to K.

Equation (16) can be expressed as

$$\mathbf{B}_k = \arg\min_{\mathbf{B}} \left\|\mathbf{X}_s - \Phi(\bar{\theta})\mathbf{B}\right\|_{\mathbf{F}}^2, \mathbf{R}_k = \mathbf{X}_s - \Phi(\bar{\theta})\mathbf{B}_k$$

and can be solved by MMVOMP [21], where the subscript k presents the k-th iteration and $\mathbf{R}_k$ represents the residual signal matrix of the k-th iteration.

D. MINIMUM NORM METHOD

In this subsection, we optimize the noise subspace $\mathbf{V}_N$ into a vector, which can improve the computing efficiency of the proposed algorithm. Theorem 1 proposes another condition to ensure that the correct DOA can be found at each iteration of OMP. $\mathbf{V}_N$ is a high matrix (i.e., the number of rows is greater than the number of columns, and the column is full rank). All noise vectors $\mathbf{V}_N(:, i), i = k+1, \ldots, N$ of the space constitute a set of complete orthogonal bases of $\mathbf{V}_N$.

Therefore, any vector $\boldsymbol{\omega} = \sum_{i=k+1}^{N} \beta_i \mathbf{V}_i$ composed of these orthogonal bases in $\mathbf{V}_N$ must also be orthogonal to $\mathbf{a}\left(\tilde{\theta}_i\right)$ because $\Delta\mathbf{R}$ and receiver noise will lead to a certain estimation error of $\mathbf{V}_N$. Based on linear algebra theory, the least norm solution $\omega_{\min}$ of $\mathbf{V}_N$ is the least sensitive to $\Delta\mathbf{V}_N$.

Solving $\boldsymbol{\omega}_{\min} = \arg\min_{\boldsymbol{\omega}\in\mathbf{F}} \boldsymbol{\omega}^H\boldsymbol{\omega}$, where $\mathbf{F} = \left\{\boldsymbol{\omega} : \boldsymbol{\omega}^T\mathbf{a}\left(\tilde{\theta}_i\right) = 0 \text{ and } \boldsymbol{\omega}(0) = 1\right\}$, it is assumed that for any vector, the following is true:

$$\boldsymbol{\omega} = \sum_{i=k+1}^{N} \beta_i \mathbf{V}_i. \tag{17}$$

Any two $\mathbf{v}_i (i = k+1, \cdots, N)$ are orthogonal, then

$$\boldsymbol{\omega}^H\boldsymbol{\omega} = \sum_{i=k+1}^{N} \beta_i^2. \tag{18}$$

The conditions of the minimum norm solution $\boldsymbol{\omega}(0) = 1$ that we are interested in can be expressed as follows:

$$\sum_{i=k+1}^{N} \beta_i \mathbf{v}_i^*(0) = 1. \tag{19}$$

Therefore, the minimum $\boldsymbol{\omega}^H\boldsymbol{\omega}$ under the constraints of $\langle \mathbf{a}_k, \boldsymbol{\omega}\rangle = 0, \boldsymbol{\omega}(0) = 1$ is equivalent to the nonconditional extreme-value for the following:

$$f(\alpha_i, \lambda) = \sum_{i=k+1}^{N} \beta_i^2 + \lambda\left(\sum_{i=k+1}^{N} \beta_i \mathbf{v}_i^*(0) - 1\right). \tag{20}$$

If:

$$\frac{\partial f}{\partial \beta} = 0, \quad i = k+1, \cdots, N. \tag{21}$$

then:

$$2\beta_i + \lambda v_i^*(0) = 0, \quad i = k+1, \cdots, N. \tag{22}$$

Substituting (22) into (17), we have

$$\boldsymbol{\omega}_{\min} = -\frac{\lambda}{2}\sum_{i=k+1}^{N} \mathbf{v}_i^*(0)\mathbf{V}_i. \tag{23}$$

From the constraints $\boldsymbol{\omega}(0) = 1$:

$$\lambda = -2\left[\sum_{i=k+1}^{MN} |\mathbf{v}_i(0)|^2\right]^{-1}. \tag{24}$$

Substituting (24) into (23), we have

$$\boldsymbol{\omega}_{\min} = \frac{\sum_{i=k+1}^{N} \mathbf{v}_i^*(0)\mathbf{v}_i}{\sum_{i=k+1}^{N} |\mathbf{v}_i(0)|^2}. \tag{25}$$

where the eigenvector $\boldsymbol{\omega}_{\min}$ is a linear combination of the orthogonal basis $\mathbf{v}_i$, with $i = k+1, \cdots, N$ in the noise vector subspace $\mathbf{V}_N$. Thus, the eigenvector $\boldsymbol{\omega}$ must be located in the noise vector subspace $\mathbf{V}_N$, and:

$$< \mathbf{a}\left(\bar{\theta}_i\right), \boldsymbol{\omega}_{\min} \geq 0. \tag{26}$$

We use (26) to reweight (6) in different angle directions to improve the performance of the proposed algorithm. The minimum norm method can improve the using efficiency of noise subspace information and reduce the weight calculation.

E. IMPLEMENTATION OF NSRomp

The NSRomp algorithm is developed by using (26) to modify the OMP algorithm. The specific process is presented as Algorithm 1.

The minimum norm method is used to combine all noise vectors in the space $\mathbf{V}_N$ into a vector, which reduces the calculations required for the NSRomp algorithm and the effect of calculation error from the noise singular vector. For M array elements and K signal sources, L snapshot estimation processing is performed. The number of atoms in the dictionary is N, and the number of steps for peak searching of the music algorithm is N. Due to the large burden number of calculations, complex multiplication is considered emphatically. With regard to the computational complexity of

TABLE 1. Simulation parameters of experiments 1-4.

Parameter	Description			
	Experiment 1	Experiment 2	Experiment 3	Experiment 4
Frequency of signal source	3.8 GHz	3.8 GHz	3.8 GHz	3.8 GHz
SNR values	5 dB	0 dB	1 dB, 10 dB	(0-10) dB
Number of snapshots	50	500	10-100,100-1000	100
Number of antenna	10	10	10	10
DOAs	$[45°, 50°, 60°]$	$[-50°, -50° + \Delta]$ $\Delta = 1° - 50°$	$[-20°, 20°, 60°]$ $[45°, 50°, 55°]$	$[45°, 50°, 60°]$
Number of runs	/	/	100	100

Algorithm 1 Noise Subspace Projection Orthogonal Matching Pursuit (NSRomp)

1) Initialization: residual $\mathbf{R}_0 = \mathbf{X}_s$, index set $\Lambda_0 = \{0\}$ and sparse representation matrix $\mathbf{\Psi}_0 = 0$.
2) At the k-th iteration:
   a) Choose the atom, which satisfies:

$$\lambda_k = \arg \max_{j=1,2...,N_s} \left\{ \frac{\|\mathbf{z}_k\|_q}{\left|\boldsymbol{\omega}_{\min}^T \mathbf{a}\left(\tilde{\theta}_j\right)\right|} \right\}. \quad (27)$$

   where $\mathbf{z}_k = \mathbf{R}_{k-1}^T \mathbf{a}\left(\tilde{\theta}_j\right)$ and $q \geq 1$.
   b) Update the index set and sparse representation matrix:

$$\mathbf{\Lambda}_k = \mathbf{\Lambda}_{k-1} \cup \{\boldsymbol{\lambda}_k\}, \mathbf{\Psi}_k = \left[\mathbf{\Psi}_{k-1}, \mathbf{a}\left(\tilde{\theta}_{\lambda_k}\right)\right]$$

   c) Reconstructing target signal: $\mathbf{B}_k = \mathbf{\Psi}_k \mathbf{X}_s$
   d) Update residual:

$$\mathbf{R}_k = \mathbf{X}_S - \mathbf{\Psi}_k^{\dagger} \mathbf{B}_k,$$

   where $\mathbf{\Psi}_k^{\dagger}$ is the pseudoinverse of $\mathbf{\Psi}_k$.

TABLE 2. Comparision of computation cost and time consumption of 3 types OMP, MUSIC and L1svd.

Experiment 1	Computational cost	Time consumption (Estimate 1 times)
MMVOMP	O(MLNK)	0.0924 s
SSVOMP	O(MNK2)+O(M3)+O(LM2)	0.0553 s
NSRomp	O[MN(K2+1)]+O(M3)+O(LM2)	0.0765 s
MUSIC	O{N[2M(M-K)+M]} +O(M3)+O(LM2)	0.9935 s
L1svd	O{(KN)3}	3.3363 s

NSRomp, the calculation of $\tilde{\mathbf{R}}_s$, its eigenvalue decomposition (EVD) in (15) requires $\mathrm{O}\left(\mathrm{LM}^2\right) + \mathrm{O}\left(\mathrm{M}^3\right)$, and solving (16) requires $\mathrm{O}\left[\mathrm{MN}\left(\mathrm{K}^2+1\right)\right]$. Table 2 shows that the new algorithm performs a moderate computational cost among the 3 types of OMP algorithms. The hardware environment used in this study includes an Intel i7-3770 CPU operating at 3.4 GHz and 16 GB of RAM.

III. SIMULATION RESULTS

In this section, we evaluate the performance of NSRomp for DOA estimation based on a 10-element uniform linear array with half-wavelength element spacing; however, the proposed algorithm is also applicable to an arbitrary array. This simulation is developed under the test condition that the DOAs of signals are just in the direction grid. Table 1 shows the simulation parameters of experiments 1-4.

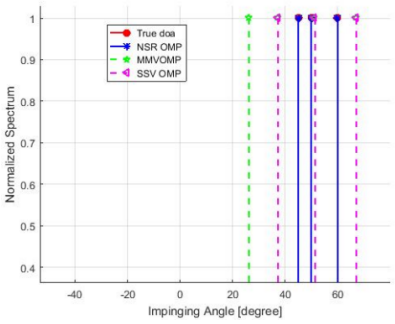


FIGURE 1. Normalized space spectra of three algorithms including MMVOMP [24], SSVOMP and proposed NSRomp.

Experiment 1 (Estimation Performance of NSRomp): In this experiment, we validate the effectiveness of the proposed NSRomp for DOA estimation and compare it with 3 types of OMPs, including MMVOMP [24], SSVOMP and propose NSRomp. To analyze the influence of the noise subspace and signal subspace on the algorithm, we call the algorithm using the signal subspace to reconstruct the signal (16) SSVOMP. When the SNR is 5 dB, the beam width of the 10-element ULA is 10.2°. Considering a set of typical DOAs(a pair of DOAs distributed in the same beam width and a DOA distributed marginally farther away), there are K = 3 uncorrelated farfield targets located at [45°, 50°, 60°]. Fig. 1 shows the normalized spatial spectrum of the three algorithms with multiple snapshots (L = 50).

The termination conditions of the three algorithms are identical: the residual energy is 20 dB below the initial signal energy, and the maximum number of iterations is M. In the Fig. 1, NSRomp and SSVOMP require three iterations to successfully reach the termination condition, while MMVOMP requires four iterations. For the [45°, 50°] signal source pairs that are relatively close, only the NSRomp yields accurate estimations; another two methods produced estimation errors. Thus, this experiment indicates that the proposed algorithm yields superior performance compared to the other two algorithms in terms of signal source resolution.

Experiment 2 (Angular Separation Experiment): In this experiment, we explore the resolution of the proposed NSRomp algorithm for DOA estimation. The angular

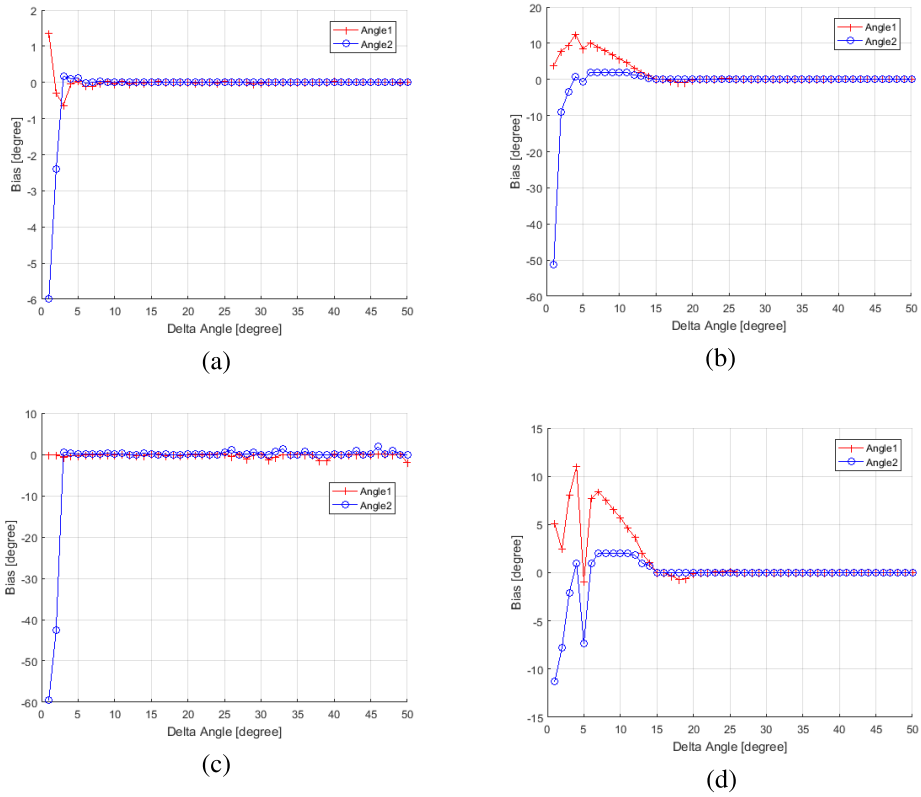


FIGURE 2. Bias versus angular separation. (a) NSRomp. (b) MMVOMP. (c) rootMUSIC. (d) SSVOMP.

separation experiment used in [34] is introduced. Assuming two equal-power uncorrelated signals arrive from Angle1 $= -50°$ and Angle2 $= -50° + \Delta$, where Δ is varied from 1° to 50° in 1° steps, the SNR is 0 dB, and the number of snapshots is 500. The bias curves of the four algorithms in Fig. 2 are obtained via 100 Monte Carlo trials and are defined on the Y-axis:

$$Bias\,(\theta_k) = \frac{1}{N}\sum_{n=1}^{\mathbf{N}}\left(\hat{\theta}_k^n - \theta_k\right), \quad k = 1, 2. \tag{28}$$

where N is the number of Monte Carlo experiments, and θ_k and $\hat{\theta}_k^n$ are the true and estimated values of the k-th DOA signal in the n-th Monte Carlo experiment, respectively.

By comparing Fig. 2 (a), (d) to (b), (c), the DOA estimation using signal subspace data, such as NSRomp and SSVOMP, exhibits a smaller angular bias when the SNR is relatively low, and the DOA of the signal is closed, which improves the estimation accuracy compared to using full-space data directly. The common feature of the NSRomp and rootMUSIC [36] algorithms is the use of noise subspace data, which tends to bias these algorithms when Δ is greater than 5°. However, the MMVOMP and SSVOMP are biased when Δ is below 15°.

By comparing Fig. 2 (a), (c) with (b), (d), the noise subspace data can alleviate the sensitivity of the algorithm to closer DOAs from signal sources.

Fig. 2 shows that the noise subspace information improves the angular resolution of the OMP algorithm from 15° to below 5°. The results of Experiment 2 show that the proposed NSRomp considers both the angular resolution and the estimation accuracy at low SNRs.

Experiment 3 (Snapshot Experiment): In this experiment, the estimation accuracy versus snapshots of the proposed algorithm is investigated and compared to MMVOMP [24], rootMUSIC [36], TLSESPRIT [37], and FOMP [38]. Two typical DOAs of signals are estimated in this experiment:

- $[-20°, 20°, 60°]$: the distance between adjacent signal sources is greater than 15°, and relative scattered signal sources are utilized;
- $[45°, 50°, 55°]$: the distance between adjacent signal sources is below one beam width, and relative close signal sources are utilized.

Each set of values is tested at a low SNR $= 1$ and a high SNR $= 10$, respectively. Because the number of snapshots required to estimate the typical value of the first group is lower, the number of snapshots increases from 10 to 100 times in intervals of 10. Also, the number of snapshots required to estimate the typical value of the second group is large. Thus, the number of snapshots increases from 100 to 1000 times in intervals of 100. 100 Monte Carlo experiments are also performed under each snapshot number, and the RMSE is

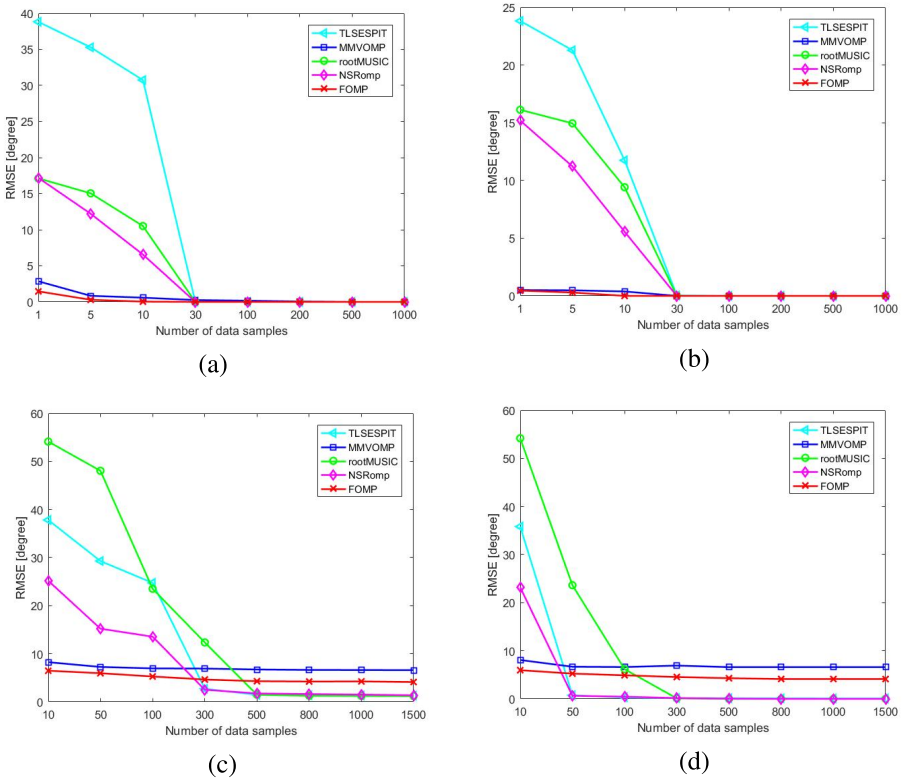


FIGURE 3. RMSE versus snapshots. (a) SNR = 1, doa=−20, 20, 60. (b) SNR = 10, doa=−20, 20, 60. (c) SNR = 1, doa=45, 50, 55. (d) SNR = 10, doa=45, 50, 55.

defined as

$$RMSE(\theta) = \sqrt{\frac{1}{NK}\sum_{k=1}^{K}\sum_{n=1}^{N}\left(\left|\hat{\theta}_k^n - \theta_k\right|^2\right)}. \tag{29}$$

where N is the number of Monte Carlo experiments, and θ_k and $\hat{\theta}_k^n$ are the true and estimated values of the k-th DOA signal in the n-th Monte Carlo experiment, respectively. Results are shown in Fig. 3.

When estimating typical values in the first group, the signal source distribution is scattered. The four types of algorithms in the experiment can all obtain high-precision estimation results within 100 snapshots. Among them, the MMVOMP algorithm does not use covariance data and does not need the perform subspace decomposition. The advantage of the MP class algorithm in the low snapshot environment is thus fully embodied. Other three algorithms that require subspace decomposition, when the snapshot number is below 30, the accuracy of the algorithm proposed in this paper is marginally higher than that of the TLSESPRIT and rootMUSIC algorithms at both low and high SNRs.

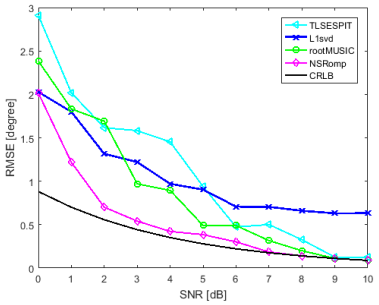
When estimating the second group of typical values, the signal source distribution is concentrated. The MMVOMP algorithm cannot calculate accurate estimations at high or low SNRs, and the results are consistent with those in experiment 2. The other three algorithms also require more snapshots. At a high SNR, their estimation accuracies are similar. When the SNR is low, and the number of snapshots is below 400, the accuracy of the proposed algorithm is marginally higher than that of the other two algorithms.

The results of Experiment 3 show that the proposed algorithm exhibits characteristics of the greedy algorithm (e.g., high estimation accuracy, even at low snapshot counts) and of subspace based algorithms (e.g., high angular resolution). Therefore, the proposed algorithm excels when the number of snapshots is low, and the signal source is concentrated.

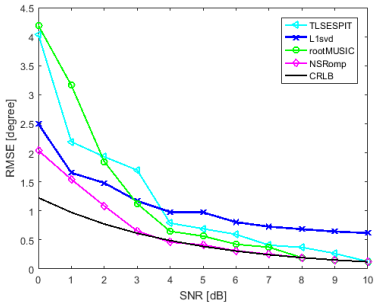
Experiment 4 (SNR Experiment): This experiment is designed to analyze the relationship between estimation accuracy and SNR. The number of snapshots L is set as 100, and the other experimental conditions are the same as those in Experiment 1. Due to the concentrated spatial distribution of signal sources, the MMVOMP and SSVOMP algorithms are invalid. Experiment 4 ignores these two algorithms and compared the performance of the proposed algorithm with rootMUSIC, TLSESPRIT, L1svd [13] and the stochastic Cramér-Rao lower bound (CRLB). 100 Monte Carlo experiments were performed for each SNR, and the results are shown in Fig. 4. The RMSE is defined as:

$$\mathrm{RMSE}\left(\theta_k\right) = \sqrt{\frac{1}{N}\sum_{n=1}^{N}\left(\left|\hat{\theta}_k^n - \theta_k\right|^2\right)}. \tag{30}$$

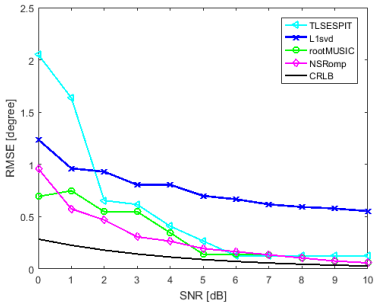
where N is the number of Monte Carlo experiments, and θ_k and $\hat{\theta}_k^n$ are the true and estimated values of the k-th DOA signal in the n-th Monte Carlo experiment, respectively.



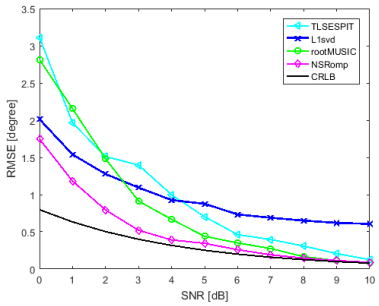
(a)



(b)



(c)



(d)

FIGURE 4. RMSE versus SNR. (a) RMSE(θ_1). (b) RMSE(θ_2). (c) RMSE(θ_3). (d) RMSE(θ).

The covariance error is large because an insufficient number of snapshots is available, and the estimation accuracy of traditional subspace algorithms is markedly affected at a low SNR. The structure of the algorithm proposed in this paper is developed based on the MP algorithm and uses the minimum norm eigenvector which is the least sensitive to errors when applying noisy subspace data. Thus, the proposed algorithm is robust to the covariance matrix estimation error. The results of experiment 4 show that by comparing (c) with (a) and (b) in Fig. 4 with θ_3, whose spatial distribution is far from the other two signal sources, the estimation accuracy of the 4 types of algorithms is significantly higher than that of θ_1, θ_2 whose relative distance is short. When the number of snapshots is relatively small, and the position of the signal source is relatively concentrated, the estimation accuracies of the two algorithms rootMUSIC and TLSESPRIT are nearly identical but marginally better than that of the L1svd algorithm. These three algorithms use only one aspect of the noise subspace or signal subspace information, while the proposed algorithm uses information from both subspaces. Experiment 4 also demonstrates that the proposed estimation algorithm yields lower MSE at most SNRs compared to existing methods.

IV. CONCLUSION

In this paper, a novel OMP algorithm for DOA estimation using subspace information is proposed. The signal subspace is used to reduce the DOA estimation problem. Compared to the multiple-snapshot OMP algorithm using full-space data (e.g., MMVOMP), the proposed algorithm reduces computation cost, improves estimation accuracy, and modifies the support set selection rule via the minimum norm eigenvector of the noise subspace. Compared to general matching pursuit algorithm that uses the highest matching degree as the support set selection criterion, the angular resolution of the algorithm is markedly improved, particularly at low SNRs. In contrast to the classic subspace algorithm, the proposed algorithm has the advantages of the MP class algorithm (e.g., short computation time and excellent performance with insufficient snapshots). The proposed algorithm also yields a better angular resolution than the MP algorithm for DOA estimation. With a concentrated signal source and few snapshots, the estimation accuracy of the proposed algorithm is higher than that of the L1svd algorithm, which requires more calculation. The proposed algorithm does not rely on the array flow pattern matrix which is valid for any array structure. Thus, it is potential to implement the MP algorithm in the field of DOA estimation in the future. The array structures in [39], [40] will be studied and compared in our future work.

REFERENCES

[1] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[2] E. J. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[3] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.

[4] Y. Fang, B. Wang, C. Sun, S. Wang, J. Hu, and Z. Song, "Joint sparsity constraint interferometric ISAR imaging for 3-D geometry of near-field targets with sub-apertures," *Sensors*, vol. 18, no. 11, p. 3750, Nov. 2018.

[5] J. Yang, X. Huang, J. Thompson, T. Jin, and Z. Zhou, "Compressed sensing radar imaging with compensation of observation position error," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 8, pp. 4608–4620, Aug. 2014.

[6] J. Yang, T. Jin, X. Huang, J. Thompson, and Z. Zhou, "Sparse MIMO array forward-looking GPR imaging based on compressed sensing in clutter environment," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 7, pp. 4480–4494, Jul. 2014.

[7] L. Wang, X. Yin, H. Yue, and J. Xiang, "A regularized weighted smoothed $L0$ norm minimization method for underdetermined blind source separation," *Sensors*, vol. 18, no. 12, p. 4260, Dec. 2018.
[8] P. Chen, D. Peng, L. Zhen, Y. Luo, and Y. Xiang, "Underdetermined blind separation by combining sparsity and independence of sources," *IEEE Access*, vol. 5, pp. 21731–21742, May 2017.
[9] G. Bao, Z. Ye, X. Xu, and Y. Zhou, "A compressed sensing approach to blind separation of speech mixture based on a two-layer sparsity model," *IEEE Trans. Audio, Speech, Language Process.*, vol. 21, no. 5, pp. 899–906, May 2013.
[10] J. Liu, K. Huang, and X. Yao, "Common-innovation subspace pursuit for distributed compressed sensing in wireless sensor networks," *IEEE Sensors J.*, vol. 19, no. 3, pp. 1091–1103, Feb. 2019.
[11] Y. Li and Y. Liang, "Compressed sensing in multi-hop large-scale wireless sensor networks based on routing topology tomography," *IEEE Access*, vol. 6, pp. 27637–27650, Sep. 2018.
[12] S. Colonnese, M. Biagi, T. Cattai, R. Cusani, F. D. V. Fallani, and G. Scarano, "Green compressive sampling reconstruction in IoT networks," *Sensors*, vol. 18, no. 8, p. 2735, Aug. 2018.
[13] 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, Aug. 2005.
[14] R. Tibshirani, "Regression shrinkage and selection via the lasso: A retrospective," *J. Roy. Stat. Soc., B (Stat. Methodol.)*, vol. 73, no. 3, pp. 273–282, Jun. 2011.
[15] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Rev.*, vol. 43, no. 1, pp. 129–159, Jan. 2001.
[16] H. Zou, "The adaptive lasso and its oracle properties," *J. Amer. Stat. Assoc.*, vol. 101, no. 476, pp. 1418–1429, Dec. 2006.
[17] B. Wang, J. Liu, and X. Sun, "Mixed sources localization based on sparse signal reconstruction," *IEEE Signal Process. Lett.*, vol. 19, no. 8, pp. 487–490, Aug. 2012.
[18] J. Fan and R. Li, "Variable selection via nonconcave penalized likelihood and its oracle properties," *J. Amer. Stat. Assoc.*, vol. 96, no. 456, pp. 1348–1360, 2001.
[19] S. Foucart and M. Lai, "Sparsest solutions of underdetermined linear systems via ℓ_q-minimization for $0 < q < 1$," *Appl. Comput. Harmon. Anal.*, vol. 26, no. 3, pp. 395–407, Mar. 2009.
[20] R. Chartrand, "Exact reconstruction of sparse signals via nonconvex minimization," *IEEE Signal Process. Lett.*, vol. 14, no. 10, pp. 707–710, Oct. 2007.
[21] S. G. Mallat, *A Wavelet Tour of Signal Processing*. Waltham, MA, USA: Academic, 1999.
[22] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3397–3415, Nov. 1993.
[23] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.
[24] J. Chen and X. Huo, "Theoretical results on sparse representations of multiple-measurement vectors," *IEEE Trans. Signal Process.*, vol. 54, no. 12, pp. 4634–4643, Dec. 2006.
[25] D. Needell and R. Vershynin, "Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 310–316, Apr. 2010.
[26] G. Karabulut, T. Kurt, and A. Yongaçoglu, "Estimation of directions of arrival by matching pursuit (EDAMP)," *EURASIP J. Wireless Commun. Netw.*, vol. 2005, no. 2, Dec. 2005, Art. no. 618605.
[27] S. Cotter, "Multiple snapshot matching pursuit for direction of arrival (DOA) estimation," in *Proc. Eur. Signal Process. Conf.*, Sep. 2007, pp. 247–251.
[28] W. Wang and R. Wu, "High resolution direction of arrival (DOA) estimation based on improved orthogonal matching pursuit (OMP) algorithm by iterative local searching," *Sensors*, vol. 13, no. 9, pp. 11167–11183, Aug. 2013.
[29] R. Roy, A. Paulraj, and T. Kailath, "ESPRIT—A subspace rotation approach to estimation of parameters of cisoids in noise," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 34, no. 5, pp. 1340–1342, Oct. 1986.
[30] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. AP-34, no. 3, pp. 276–280, Mar. 1986.
[31] C. Zheng, G. Li, H. Zhang, and X. Wang, "An approach of DOA estimation using noise subspace weighted $\ell 1$ minimization," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2011, pp. 2856–2859.
[32] X. Xu, X. Wei, and Z. Ye, "DOA estimation based on sparse signal recovery utilizing weighted l_1-norm penalty," *IEEE Signal Process. Lett.*, vol. 19, no. 3, pp. 155–158, Mar. 2012.
[33] N. Hu, Z. Ye, D. Xu, and S. Cao, "A sparse recovery algorithm for DOA estimation using weighted subspace fitting," *Signal Process.*, vol. 92, no. 10, pp. 2566–2570, Oct. 2012.
[34] J. Yin and T. Chen, "Direction-of-arrival estimation using a sparse representation of array covariance vectors," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4489–4493, Sep. 2011.
[35] 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, Feb. 2011.
[36] B. D. Rao and K. V. S. Hari, "Performance analysis of root-MUSIC," *IEEE Trans. Acoust., Speech Signal Process.*, vol. 37, no. 12, pp. 1939–1949, Dec. 1989.
[37] R. Roy and T. Kailath, "Total least squares ESPRIT," in *Proc. 21st Asilomar Conf. Signals, Syst., Comput.*, Monterey, CA, USA, Nov. 1987, pp. 297–301.
[38] M. Dehghani and K. Aghababaiyan, "FOMP algorithm for direction of arrival estimation," *Phys. Commun.*, vol. 26, pp. 170–174, Feb. 2018.
[39] K. Aghababaiyan, R. G. Zefreh, and V. Shah-Mansouri, "3D-OMP and 3D-FOMP algorithms for DOA estimation," *Phys. Commun.*, vol. 31, pp. 87–95, Dec. 2018.
[40] K. Aghababaiyan, V. Shah-Mansouri, and B. Maham, "High-precision OMP-based direction of arrival estimation scheme for hybrid non-uniform array," *IEEE Commun. Lett.*, vol. 24, no. 2, pp. 354–357, Feb. 2020.

YANG ZHAO received the B.Eng. degree from Jilin University, China, in 2009. He is currently pursuing the Ph.D. degree in science with Jilin University. His research interests include compressed sensing and DOA tracking.

SI QIN received the B.Eng. and M.Sc. degrees from Jilin University, in 2009 and 2012, respectively. She is currently an Assistant Professor with the Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences. Her research interests include signal processing and image processing.

YI-RAN SHI (Member, IEEE) received the B.Eng. degree from the Changchun University of Science and Technology, China, in 2007, the M.Sc. degree from the Lanzhou University of Technology, China, in 2010, and the Ph.D. degree from Jilin University, China, in 2014, all in control engineering. He is currently an Associate Professor with Jilin University. He is the author or the coauthor of more than 30 articles. His current research interests include signal processing and model predictive control.

YAO-WU SHI received the B.Eng., M.Sc., and Ph.D. degrees from the Jilin University of Technology, China, in 1982, 1985, and 1994, respectively. From 1982 to 1993, he was a Lecturer with the Jilin University of Technology, where he was promoted to an Associate Professor, in 1998, and to a Professor of Jilin University, in 2003. He is the author or the coauthor of one book, several book chapters, and more than 200 articles. His current research interests include modern cross spectrum analysis, system identification, DOA estimation, harmonic parameter estimation, and chaotic signal processing.