Superresolution 2D DOA Estimation for a Rectangular Array via Reweighted Decoupled Atomic Norm Minimization

Ming-Ming Liu, Chun-Xi Dong, Yang-Yang Dong, and Guo-Qing Zhao

School of Electronic Engineering, Xi’an Jiaotong University, Xi’an 710071, China

Correspondence should be addressed to Ming-Ming Liu; 18710981090@163.com

Received 13 April 2019; Revised 1 June 2019; Accepted 19 June 2019; Published 8 July 2019

This paper proposes a superresolution two-dimensional (2D) direction of arrival (DOA) estimation algorithm for a rectangular array based on the optimization of the atomic \( \ell_0 \) norm and a series of relaxation formulations. The atomic \( \ell_0 \) norm of the array response describes the minimum number of sources, which is derived from the atomic norm minimization (ANM) problem. However, the resolution is restricted and high computational complexity is incurred by using ANM for 2D angle estimation. Although an improved algorithm named decoupled atomic norm minimization (DAM) has a reduced computational burden, the resolution is still relatively low in terms of angle estimation. To overcome these limitations, we propose the direct minimization of the atomic \( \ell_0 \) norm, which is demonstrated to be equivalent to a decoupled rank optimization problem in the positive semidefinite (PSD) form. Our goal is to solve this rank minimization problem and recover two decoupled Toeplitz matrices in which the azimuth-elevation angles of interest are encoded. Since rank minimization is an NP-hard problem, a novel sparse surrogate function is further proposed to effectively approximate the two decoupled rank functions. Then, the new optimization problem obtained through the above relaxation can be implemented via the majorization-minimization (MM) method. The proposed algorithm offers greatly improved resolution while maintaining the same computational complexity as the DAM algorithm. Moreover, it is possible to use a single snapshot for angle estimation without prior information on the number of sources, and the algorithm is robust to noise due to its iterative nature. In addition, the proposed surrogate function can achieve local convergence faster than existing functions.

1. Introduction

The problem of two-dimensional (2D) direction of arrival (DOA) estimation is encountered in various signal processing applications, such as 2D angle estimation using 2D arrays in multiple-input and multiple-output (MIMO) communication systems. Although numerous algorithms for this purpose have emerged in the past three decades [1–5], there is still an urgent need for fast and superresolution algorithms that can utilize bits of snapshots, especially for massive systems.

Conventional superresolution algorithms such as MUSIC and ESPRIT [6–9] are based on subspace techniques. Considering that these methods either require a heavy computational load for the 2D spectrum search or are suitable only for arrays that exhibit geometric invariance based on a number of samples that is larger than the number of sensors, we instead focus on algorithms that rely on the popular theory of compressed sensing (CS), which makes it possible to use a smaller number of snapshots, or even a single snapshot, to estimate the DOA, even when the sources are correlated. However, another challenge arises due to grid mismatch [10, 11].

Recently, the mathematical concept of continuous compressed sensing (CCS) was introduced by Candès and Fernandez-Granda [12]. CCS no longer follows traditional grid-based models, which require linear independence among the grids, but rather relies on a superresolution concept based on the total variation norm; thus, it fundamentally avoids the problem of grid mismatch and further facilitates the application of CS in parameter estimation. Subsequently, Chandrasekaran et al. [13] and Tang et al. [14] developed an atomic norm metric and presented an atomic norm minimization (ANM) problem in the positive semidefinite (PSD) form utilizing the Vandermonde structure of the atoms. The theory was further developed for the estimation of one-dimensional (1D) parameters by Yang et al. [15], who
also proposed the reweighted atomic norm minimization (RAM) algorithm to enhance the resolution [16]. In addition, the metric was extended for higher-dimensional parameter estimation by Xu et al. [17] and the authors [18, 19]. However, the computational complexity is enormous even for 2D applications. To combat this issue, Tian et al. [20] presented a decoupled atomic norm minimization (DAM) algorithm in which the two-level Toeplitz matrix is divided into two one-level matrices, thereby improving the calculation efficiency. However, the resolution is limited because the minimum angular interval is inversely proportional to the number of rows or columns in the array [12, 19, 20].

To address the above issues, we propose a low-complexity superresolution optimization algorithm for 2D DOA estimation with a rectangular array. In essence, the proposed algorithm adopts an extended atomic $\ell_0$ norm that avoids the constraint on the angular intervals, and we argue that the problem can be translated into a decoupled rank minimization problem in the PSD form, whereafter since the rank minimization problem is difficult to solve directly, a novel sparse surrogate for it is presented, inspired by the works [16, 21]. Then, the new concave objective function can be iteratively optimized by the well-known majorization-minimization (MM) method. Our algorithm eventually takes the form of a weighted DAM problem in each iteration; thus, we name it reweighted decoupled atomic norm minimization (RDAM). A series of proofs are presented in this paper to illustrate the rationality of the proposed method, and corresponding simulations are also provided to show its validity. Our proposed algorithm enhances the resolution to a great extent without increasing the computational burden, and it requires neither prior information on the number of sources nor the noise level which can be obtained in the iterative process. In addition, our method is more robust to noise than the DAM algorithm due to its iterative nature.

2. Signal Model

Suppose that $K$ far-field narrowband sources $\{s_k(t)\}_{k=1}^K$ are impinging on a uniform rectangular array (URA) composed of $N \times M$ sensors, where $N$ and $M$ are the numbers of elements in the $x$ and $z$ directions, respectively. For the sake of subsequent discussion, we adopt a decoupled observation model for the array output [22]. As shown in Figure 1, we redefine the azimuth $\theta$ of the received signal as the angle between the signal and the $yoz$ plane instead of using the traditional definition; the definition of the elevation $\phi$ remains unchanged. Correspondingly, the single-snapshot array observation model without noise is

$$X^o = \sum_{k=1}^K s_k a_N(\theta_k) a_N^H(\phi_k) \tag{1}$$

where $X^o$ is the array response, with dimensions of $N \times M$, and $s_k = |s_k|e^{j\theta_k}$ is the $k$-th received source, with $|s_k|$ denoting the amplitude of the source and $\theta_k$ denoting the phase, which takes values in the range $[0, 2\pi]$ (the frequency part of each source is ignored for simplicity). Moreover, $a_N(\theta_k) = [\xi_{1k}, \xi_{2k}, \ldots, \xi_{MK}]^T$ and $a_N(\phi_k) = [\eta_{1k}, \eta_{2k}, \ldots, \eta_{MK}]^T$ are the 1D steering vectors in the $x$ and $z$ directions, respectively. Suppose that the array spacing is set to half the wavelength; then, we have $\xi_{nk} = \exp(-j2\pi \cdot 0.5(n-1)\sin\theta_k)$ and $\eta_{nk} = \exp(j2\pi \cdot 0.5(n-1)\sin\phi_k)$. Note that $\eta_{nk}$ is the conjugate of the regular vector.

Our goal in this paper is to estimate $\theta_k$ and $\phi_k$ from the observed data $X^o$, and we assume that the following conditions hold:

1. $a_N(\theta_k)$ and $a_N(\phi_k)$ have the Vandermonde structure.
2. $0.5 \sin \theta_k \in [-1/2, 1/2]$ and $0.5 \sin \phi_k \in [-1/2, 1/2]$.
3. The phases $2[\theta_k]_{k=1}^K$ are either i.i.d. and uniformly drawn from the complex unit circle or independent random variables with mean $E[\theta_k] = 0$.

The first two conditions, which obviously hold for array models, are the basic requirements for the algorithm subsequently proposed in this paper, and the last one guarantees that the solution of our proposed algorithm is exact with a probability of at least $1 - \delta$, where $\delta$ is a very small value [14, 19, 23].

3. Proposed Reweighted Decoupled Atomic Norm Minimization (RDAM) Algorithm

3.1. Decoupled Rank Minimization Problem. The atomic $\ell_0$ norm was proposed in [14] as a measure of the sparsest number of spikes with unknown positions that could compose the observed signal in the frequency domain. In this paper, we employ this norm to express the number of sources with unknown directions composing the array output in the spatial domain and implement 2D DOA estimation by optimizing the atomic representation.

More precisely, we merge the signal phase information into the steering vectors and redefine them as follows:

$$a_N\left(\theta_k, \frac{1}{2} \theta_k\right) = \exp\left(-j2\pi \cdot 0.5(n-1)\sin\theta_k + \frac{1}{2} \theta_k\right), \tag{2}$$

$$n = 1, \ldots, N$$

![Figure 1: Signal model.](https://example.com/figure1.png)
Proposition 1. Let \( X^0 \in C^{N \times M} \) be an array observed data that has the following atomic form and satisfies the conditions given in the previous section:

\[
X^0 = \sum_{k=1}^{K} |s_k| A(\theta_k, \varphi_k, \delta_k)
\]

(6)

The corresponding atomic \( \ell_0 \) norm minimization problem can be expressed as

\[
\min_X \|X\|_{\ell_0, 0} \quad \text{s.t.} \quad X = X^0
\]

(7)

Then, (7) is equivalent to the optimum value of the following decoupled rank minimization problem:

\[
\min_{u_0, u_p} X \quad \frac{1}{2} \text{rank}(T(u_0)) + \frac{1}{2} \text{rank}(T(u_p))
\]

s.t.

\[
\begin{bmatrix} T(u_0) & X^H \end{bmatrix} \begin{bmatrix} X & T(u_0) \end{bmatrix} \succeq 0, \quad X = X^0
\]

(8)

where \( T(u_0) \) and \( T(u_p) \) denote two 1D Toeplitz matrices and \( u_0 \) and \( u_p \) are their first rows, respectively.

Proof. The trivial case of \( X = 0 \) is unnecessary to discuss. For \( X \neq 0 \), let \( r_1^* \) and \( r_2^* \) be the optimal values of \( \text{rank}(T(u_0)) \) and \( \text{rank}(T(u_p)) \), respectively; then, the corresponding optimal value of (8) can be expressed as \( (1/2)r_1^* + (1/2)r_2^* \). Furthermore, let the optimal \( \|X\|_{\ell_0, 0} \) be denoted by \( \mathcal{X}^* \).

On one side, we will show that \( (1/2)r_1^* + (1/2)r_2^* \leq \mathcal{X}^* \) given problem (7). Suppose that \( \|X\|_{\ell_0, 0} = \mathcal{X}^* < \min[N, M] \) can be realized in the arbitrary atomic decomposed form \( X = \sum_{k=1}^{K^*} |s_k| a_M(\theta_k, \varphi_k, \delta_k) \), where \( \{|s_k|\}_{k=1}^{K^*} \) is a set of positive real numbers. Let \( u_0 = \sum_k |s_k| a_M(\theta_k, 0) \) and \( u_p = \sum_k |s_k| a_M(\varphi_k, 0) \); then, the homologous Toeplitz matrices can be represented as

\[
T(u_0) = \sum_k |s_k| a_M(\theta_k, 0) a_M^H(\theta_k, 0)
\]

(9)

and

\[
T(u_p) = \sum_k |s_k| a_M(\varphi_k, 0) a_M^H(\varphi_k, 0)
\]

(10)

Thus, we obtain the following PSD matrix:

\[
\begin{bmatrix} T(u_0) & X^H \\ X & T(u_0) \end{bmatrix} = \sum_{k=1}^{K^*} \begin{bmatrix} a_M(\theta_k, \varphi_k, \delta_k) & 0 \\ 0 & a_M^H(\theta_k, \varphi_k, \delta_k) \end{bmatrix} \succeq 0
\]

(11)

i.e., the relational expressions \( r_1^* \leq \mathcal{X}^* \) and \( r_2^* \leq \mathcal{X}^* \) hold. Hence, \( (1/2)r_1^* + (1/2)r_2^* \leq \mathcal{X}^* \).

On the other side, given problem (8), we will try to verify that \( K^* \leq (1/2)r_1^* + (1/2)r_2^* \) in the case of \( r_1^* < N \) and \( r_2^* < M \) (the full-rank case is trivial). Given the PSD condition, and letting \( u_0^* \) and \( u_p^* \) compose the optimal solution to (8), the following can be derived in accordance with the Schur Complement Lemma [27, 28]:

\[
\begin{bmatrix} T(u_0^*) & X^H \\ X & T(u_0^*) \end{bmatrix} \succeq 0
\]

(12)
optimize the decoupled rank minimization problem (8)

\[ T(\mathbf{u}_p^*) \mathbf{X} 
\begin{bmatrix} 
\mathbf{X}^H 
T(\mathbf{u}_p^*) 
\end{bmatrix} \succeq 0 
\]

implies that the function \( T(\mathbf{u}_p^*) \mathbf{X} \) is positive semidefinite, which means that \( \mathbf{X} \) can be decomposed as \( \mathbf{X} = \mathbf{V} \mathbf{D} \mathbf{V}^H \), where \( \mathbf{V} \) is a unitary matrix and \( \mathbf{D} \) is a diagonal matrix with positive real elements. Therefore, (12) implies that \( \mathbf{X} \) lies in the range of \( \mathbf{V} \), which means that \( \mathbf{X} \) can be composed of no more than \( r_1^* \) atoms; i.e., \( \mathcal{X}^* \leq r_1^* \). Similarly, (13) leads to the conclusion that \( \mathcal{X}^* \leq r_2^* \).

Based on the above proposition, we can straightforwardly optimize the decoupled rank minimization problem (8) under the PSD constraint to implement the optimization of (7). It is evident from the above proof that \( \theta_k^* \) and \( \varphi_k^* \) are encoded in the Toeplitz matrices \( T(\mathbf{u}_p^*) \) and \( T(\mathbf{u}_q^*) \), respectively. Once the optimal solution \( (\mathbf{u}_p^*, \mathbf{u}_q^*, \mathbf{X}^*) \) is obtained, \( \theta_k^* \) and \( \varphi_k^* \) can be determined from the Vandermonde decompositions of these two matrices; there are many off-the-shelf computational methods such as the dual root-finding method [12, 14, 15] and the matrix pencil method [29, 30]. In addition, supposing \( \theta_k^* \) and \( \varphi_k^* \) are acquired, the simple and straightforward method of angle pairing provided in [20] can be applied. Therefore, our task is simply to optimize problem (8) for \( (\mathbf{u}_p^*, \mathbf{u}_q^*, \mathbf{X}^*) \).

3.2. A Novel Sparse Surrogate for Rank Minimization. The essential objective of the decoupled rank minimization problem (8) discussed above is to recover two low-rank matrices within the PSD feasible set. Unfortunately, due to the difficulty of solving this problem (which is similar to that of solving the original optimization problem), there is an urgent need to seek new metrics to express the objective function. The main idea is to find one simpler surrogate function that not only can effectively simulate the ranks of the two Toeplitz matrices but also can simultaneously encourage a low-rank solution. In most papers that discuss the rank minimization problem [16, 27, 31, 32], the log-det function \( f_1(T(\mathbf{u})) \), which is expressed as follows, is generally used:

\[ f_1(T(\mathbf{u})) = \ln |T(\mathbf{u}) + \epsilon_1 I| = \ln \left( \frac{1}{\epsilon_1} T(\mathbf{u}) + I \right) + \ln \epsilon_1 \]

However, in this paper, we consider a novel concave surrogate function for \( T(\mathbf{u}) \in \mathbb{C}^{Q \times Q} \), given by

\[ f_2(T(\mathbf{u})) = -\epsilon_2 \ln \left( (T(\mathbf{u}) + \epsilon_2 I)^{-1} \right) \]

where \( I \) represents the identity matrix. The regularization parameter \( \epsilon_2 > 0 \) enables the function \( f_2(T(\mathbf{u})) \) to approximate \( \text{rank}(T(\mathbf{u})) \) as it varies and prevents the function from taking a value of \(-\infty\) when \( T(\mathbf{u}) \) is rank-deficient. Moreover, note that the parameter \( \epsilon_1 \) does the same job with \( f_1(T(\mathbf{u})) \).

From the properties of matrix eigenvalues and trace, it follows that

\[ \epsilon_2 \left( (T(\mathbf{u}) + \epsilon_2 I)^{-1} \right) = \frac{1}{\lambda_i + \epsilon_2} \]

for any \( i \),

\[ f_2(T(\mathbf{u})) = -\sum \frac{\epsilon_2}{\lambda_i + \epsilon_2} = -\sum \frac{1}{\lambda_i/\epsilon_2 + 1} \]

where \( \lambda^Q_{i=1} \) is the set of eigenvalues of \( T(\mathbf{u}) \). Equation (17) implies that the essence of function \( f_2(T(\mathbf{u})) \) is to operate on \( \left\{ \lambda_i \right\}_{i=1}^Q \); thus, to better understand the rank-approximation property of \( f_2(T(\mathbf{u})) \) on a PSD definition domain, we denote a new subfunction by

\[ p_2(\lambda_i) = -\left( \frac{\lambda_i}{\epsilon_2} + 1 \right)^{-1} + 1 \]

for any eigenvalue \( \lambda_i \geq 0 \), which is an inverse function obtained through a certain translation and scaling process and has the features of monotonicity and concavity. Note that the constant term 1 is added merely for consistency with the \( \| \cdot \|_{\ell_0} \) norm; it is equivalent to adding an identity matrix to \( f_2(T(\mathbf{u})) \) and does not influence the outcome. Obviously, we have the identity \( p_2(0) = 0 \); at \( \lambda_i > 0 \) and near the zero point, \( p_2(\lambda_i) \) has an increasingly steep slope of \( 1/\epsilon_2 \) as \( \epsilon_2 \to 0 \), which equivalently imposes a large penalty on near-zero \( \lambda_i \) and ensures that the values of \( p_2(\lambda_i) \) at these \( \lambda_i \) become much smaller.

In addition, \( p_2(\lambda_i) \) tends towards a constant value of 1 for a sufficiently large \( \lambda_i \). The above trend illustrates that as \( \epsilon_2 \to 0 \), \( p_2(\lambda_i) \) favours the separation of \( \lambda_i \) at different levels, thus becoming a relaxation of the \( \| \cdot \|_{\ell_0} \) norm \( \| \cdot \|_{\ell_0} \), which has an infinite slope near the zero point and a constant value of 1 except at the zero point. To provide an intuitive illustration, we plot \( p_2(\lambda_i) \) along different \( \epsilon_2 \)‘s at \( \lambda_i \geq 0 \) in Figure 2.
Meanwhile, the curves of the $\ell_0$ norm, the $\ell_1$ norm, and the log-det subfunction represented by $p_1(\lambda_j) = \ln(\lambda_j / \epsilon_1 + 1)$ are provided for comparison. Suitable proportionality constants are set to ensure that $p_2(\lambda_j)$ and $p_1(\lambda_j)$ are zero at $\lambda_j = 0$ and have a constant value of 1 at $\lambda_j = 1$ without modifying their properties. As shown in Figure 2, both $p_2$ and $p_1$ approach the $\ell_0$ norm as the regularization parameters decrease; the smaller $\epsilon_2$ is, the steeper the slope is, which offers a basis for the selection of $\epsilon_2$ during both initialization and iteration, as discussed in Section 3.3. Moreover, the function $p_2$ is more likely to perform better since it has a much faster speed of approximation; i.e., $p_2$ is deemed to be more sparsity-enhancing.

The approximate relationship between $p_2(\lambda_j)$ and $\|\lambda\|_{\ell_0}$ provides an explanation for why rank($T(u)$) can be relaxed by $f_2(T(u))$. In particular, since each larger $\lambda_j$ corresponds to a value of $p_2(\lambda_j)$ that tends towards a constant value of 1 as $\epsilon_2 \rightarrow 0$ and the addition operation does not affect the monotonicity and concavity of a function, the sum of all $p_2(\lambda_j)$ becomes a relaxation of the number of larger eigenvalues, which is considered to be the rank of the Hermitian matrix $T(u)$ or the number of sources; i.e., the sparse surrogate $f_2(T(u))$ can be regarded as a relaxation of rank($T(u)$). Furthermore, the proposed function $f_2(T(u))$ leads to better performance than $f_1(T(u))$ in encouraging low-rank solutions, as seen by comparing the curves for $p_2$ with the curves for $p_1$; stronger interpretations will be presented in later subsections.

3.3. RDAM via the Majorization-Minimization (MM) Algorithm. Using the surrogate function $f_2(T(u))$, the decoupled rank minimization problem (8) can be relaxed as follows:

$$
\begin{align*}
\min_{u, u_0, X} & \quad -\frac{1}{2} \epsilon_0 \text{tr} \left((T(u_0) + \epsilon_0 I)^{-1}\right) \\
& \quad -\frac{1}{2} \epsilon_0 \text{tr} \left((T(u) + \epsilon_0 I)^{-1}\right) \\
\text{s.t.} & \quad \begin{bmatrix} T(u_0) & X^H \\ X & T(u) \end{bmatrix} \succeq 0, \quad X = X^c \end{align*}
$$

(18)

Here, the parameters $\epsilon_0$ and $\epsilon_0$ are deemed to be mutually independent. In this subsection, we aim to minimize the objective function in (18), denoted by $F(T(u_0), T(u))$, on the PSD cone $PSD(u_0, u)$. Considering that $F(T(u_0), T(u))$ is a sum of two concave functions, the MM algorithm [33, 34] can be adopted to obtain a local optimum. This algorithm follows an iterative procedure consisting of two steps, i.e., majorization and minimization, in each iteration.

Given an initial point $\{T(u_{0,0}), T(u_{0,0})\} \in PSD(u_0, u)$, a series of feasible points $\{T(u_{0,j}), T(u_{0,j})\}$, $j = 1, 2, \ldots$ are generated. At each feasible point, the majorization step requires the construction of a tractable function majorizing $F(T(u_0), T(u))$ while ensuring that its value at the current point $\{T(u_{0,j}), T(u_{0,j})\}$ is equal to $F(T(u_{0,j}), T(u_{0,j}))$. Here, we use the first-order Taylor expansion for this purpose. Since both terms of $F(T(u_0), T(u))$ possess the same structure, we consider only the first term as an example, for which we have the following first-order approximate expression:

$$
\epsilon_0 \text{tr} \left( (T(u_{0,j}) + \epsilon_0 I)^{-1} \right) + \epsilon_0 \text{tr} \left( (T(u_{0,j}) + \epsilon_0 I)^{-2} (T(u_0) - T(u_{0,j})) \right) = \epsilon_0 \text{tr} \left( (T(u_{0,j}) + \epsilon_0 I)^{-2} (T(u_0)) + \epsilon_0 \right)
$$

(19)

where $\epsilon_0$ is a constant. The derivation is presented in Appendix B. Thus, $F(T(u_0), T(u))$ has the following upper bound:

$$
\begin{align*}
& F \left( T(u_0), T(u) \right) \\
& \leq \frac{\epsilon_0}{2} \text{tr} \left( (T(u_{0,j}) + \epsilon_0 I)^{-2} (T(u_0)) \right) \\
& \quad + \frac{\epsilon_0}{2} \text{tr} \left( (T(u_{0,j}) + \epsilon_0 I)^{-2} (T(u_0)) \right) + C_j \\
& \overset{\Delta}{=} G \left( \{T(u_0), T(u)\} : \{T(u_{0,j}), T(u_{0,j})\} \right)
\end{align*}
$$

(20)

where $C_j = \epsilon_0 + \epsilon_0$ is constant. It is evident that $G(T(u_0), T(u)) : \{T(u_{0,j}), T(u_{0,j})\}) = F(T(u_{0,j}), T(u_{0,j}))$. In addition, many other construction methods [33] (e.g., second-order Taylor expansion, inequality, and the Schur complement) could be used; these possibilities are left for future work.

Then, in the minimization step, the following minimization problem is solved, ignoring the constant $C_j$, to update $T(u_{0,j+1})$ and $T(u_{0,j+1})$:

$$
\begin{align*}
\min_{u_{0,j+1}, u_{0,j+1}, X_{j+1}} & \quad \frac{1}{2} \text{tr} (\epsilon_0 (T(u_{0,j}) + \epsilon_0 I)^{-2} (T(u_0)) + \frac{1}{2} \text{tr} (\epsilon_0 (T(u_{0,j}) + \epsilon_0 I)^{-2} (T(u_0)) \\
\text{s.t.} & \quad \begin{bmatrix} T(u_{0,j}) & X^H \\ X & T(u_0) \end{bmatrix} \succeq 0, \quad X = X_j
\end{align*}
$$

(21)

where the optimization problem is a semidefinite programming (SDP) that can be solved in polynomial time. Notice that this problem is equivalent to a weighted DAM problem in the PSD form [20], with weighting factors $W_{0,j} = \epsilon_0 (T(u_{0,j}) + \epsilon_0 I)^{-2}$ and $W_{0,j} = \epsilon_0 (T(u_{0,j}) + \epsilon_0 I)^{-2}$, and that the proposed algorithm is iterative in nature; therefore, our
algorithm is called RDAM. In addition, we have the following inequality:

\[
F \left( T \left( \mathbf{u}_{0,j+1} \right), T \left( \mathbf{u}_{p,j+1} \right) \right) \\
\leq G \left( \{ T \left( \mathbf{u}_{0,j} \right), T \left( \mathbf{u}_{p,j} \right) \} : \{ T \left( \mathbf{u}_{0,j} \right), T \left( \mathbf{u}_{p,j} \right) \} \right) \\
+ C_j
\]

\[
\leq G \left( \{ T \left( \mathbf{u}_{0,j} \right), T \left( \mathbf{u}_{p,j} \right) \} : \{ T \left( \mathbf{u}_{0,j} \right), T \left( \mathbf{u}_{p,j} \right) \} \right) \\
+ C_j = F \left( T \left( \mathbf{u}_{0,j} \right), T \left( \mathbf{u}_{p,j} \right) \right);
\]

i.e., \( \{ F(T(\mathbf{u}_{0,j})), T(\mathbf{u}_{p,j})) \}_{j \in J} \) is a nonincreasing sequence, implying that the iterative process makes the objective function \( F(\mathbf{T}(\mathbf{u}_{0})), \mathbf{T}(\mathbf{u}_{p})) \) repeatedly smaller, and it can be expected to converge to a local minimum \( F(\mathbf{T}(\mathbf{u}_{0}^{*})), \mathbf{T}(\mathbf{u}_{p}^{*})) \), as proven in detail in [21, 33]. Furthermore, since the algorithm is locally convergent, suitable initial values are indispensable for guaranteeing convergence to the desired local optimum. One immediate possibility is to use the DAM solution for initialization. In particular, let \( \mathbf{W}_{\theta,0} \) and \( \mathbf{W}_{\varphi,0} \) be identity matrices; then, we can obtain the initial values \( T(\mathbf{u}_{0,j}) \) and \( T(\mathbf{u}_{p,j}) \) by solving a DAM problem. Next, let us emphasize the selection of \( \varepsilon_{\theta,j} \) and \( \varepsilon_{\varphi,j} \), both of which influence the outcome. As argued in Section 3.2, \( F(T(\mathbf{u}_{0}), T(\mathbf{u}_{p})) \) is the sum of two subfunctions, denoted by \( p_{2}(\lambda_{\theta,j}) = -(\lambda_{\theta,j}/\varepsilon_{\theta,j} + 1)^{-1} \) and \( p_{2}(\lambda_{\varphi,j}) = -(\lambda_{\varphi,j}/\varepsilon_{\varphi,j} + 1)^{-1} \), for \( \theta \) and \( \varphi \), respectively. For consistency with the curves in Figure 2, we set \( \varepsilon_{\theta,j} = a_{\lambda_{\theta,max}} \) and \( \varepsilon_{\varphi,j} = b_{\lambda_{\varphi,max}} \). Then, both the factors \( a \) and \( b \) can be viewed as equivalent to the parameter \( \varepsilon_{\lambda} \) in this figure, and we have \( \lambda_{\theta,j}/\lambda_{\theta,max} \in [0, 1] \) and \( \lambda_{\varphi,j}/\lambda_{\varphi,max} \in [0, 1] \). Now, let us consider the choice of \( a \) as an example. Larger eigenvalues, corresponding to the signals, are denoted by \( \lambda_{\theta,max} \); smaller eigenvalues, corresponding to the noise, are denoted by \( \lambda_{\theta,max} \); thus, we should be able to ensure a relatively large difference in the values of \( p_{2}(\lambda_{\theta,j}) \) between the points \( \lambda_{\theta,j}/\lambda_{\theta,max} \) and \( \lambda_{\theta,max}/\lambda_{\theta,max} \) by tuning the initial factor \( a \). For example, for points \( \lambda_{\theta,j}/\lambda_{\theta,max} \) and \( \lambda_{\theta,max}/\lambda_{\theta,max} \) that are close to each other, we can select a small \( a \) value, such as 0.1 or smaller; then, there will be a sharp slope such that the subfunction values will be distinctly separated. In contrast, for the relatively distant points, a large \( a \) value, such as 1, will be sufficient to achieve good estimation performance. The same analysis applies for the initialization of the factor \( b \).

Finally, it is instructive to explore the potential implications of the RDAM algorithm in combination with the atomic norm [13, 14]. Notice that the concept of the 1D weighted atom set proposed in [16] by Yang et al. can be extended to a new 2D weighted atom set \( \mathcal{B}^{-} \) as follows:

\[
\mathcal{B}^{-} = \left\{ w_{\theta} \left( \theta \right) a_{N} \left( \theta, \frac{1}{2} \theta \right) w_{\varphi} \left( \varphi \right) a_{M} \left( \varphi, \frac{1}{2} \theta \right) : \mathcal{B}^{-} \right\}
\]

where \( \{ w_{\theta}(\theta), w_{\varphi}(\varphi) \} \) denotes a pair of nonnegative weights on the 2D space angles. Accordingly, the weighted atom norm is given by

\[
\| \mathbf{X} \|_{\mathcal{B}^{-}}^{-} = \inf \{ t > 0 : \mathbf{X} \in t \mathbf{conv} (\mathcal{A}^{-}) \} = \inf \left\{ \sum_{k} |s_{k}| : \mathbf{X} = \sum_{k} |s_{k}| \right\}.
\]

Then, a 2D weighted ANM problem in decoupled PSD form can be obtained based on the following proposition.

**Proposition 2.** Let \( \| \mathbf{X} \|_{\mathcal{B}^{-},j+1} \) denote the \((j+1)\)th 2D weighted atomic norm. The following equivalence relation holds given the data \( \mathbf{X}_{j} \in C^{N \times M} \) recovered from the \( j \)th iteration.

\[
\min \| \mathbf{X} \|_{\mathcal{B}^{-},j+1} = \min_{\mathbf{u}_{0,j+1} \in \mathbf{u}_{0,j+1}} \left\{ \frac{1}{2} \text{tr} \left( \mathbf{W}_{\theta,j} \mathbf{T}(\mathbf{u}_{0}) \right) + \frac{1}{2} \text{tr} \left( \mathbf{W}_{\varphi,j} \mathbf{T}(\mathbf{u}_{p}) \right) \right\}
\]

\[
s.t. \left[ \mathbf{T}(\mathbf{u}_{0}), \mathbf{X}^{H} \right] \mathbf{X}^{H} \mathbf{T}(\mathbf{u}_{0}) \geq 0, \quad \mathbf{X} = \mathbf{X}_{j},
\]

where

\[
\omega_{\theta,j}(\theta) = \frac{1}{\sqrt{a_{N}^{H}(\theta) \mathbf{W}_{\theta,j} a_{N}(\theta)}}
\]

\[
\omega_{\varphi,j}(\varphi) = \frac{1}{\sqrt{a_{M}^{H}(\varphi) \mathbf{W}_{\varphi,j} a_{M}(\varphi)}}.
\]

This implies that iteratively minimizing the weighted atomic norm \( \| \mathbf{X} \|_{\mathcal{B}^{-}} \) induced by the 2D weighted atom set \( \mathcal{B}^{-} \) serves the same purpose as the proposed RDAM algorithm; i.e., RDAM can also be called decoupled reweighted atom norm minimization (DRAM). More significantly, our proposed trace function \( f_{2}(\mathbf{T}(\mathbf{u})) \) gives rise to quadratic weighting factors, which greatly accelerates the speed of local convergence and further enhances the preference for low-rank solutions in (18) compared with the first-order factors generated by the log-det function \( f_{1}(\mathbf{T}(\mathbf{u})) \) [16].
3.4. Another Interpretation from the Perspective of Conventional Direction of Arrival (DOA) Estimation Methods. The proposed RDAM algorithm can also be interpreted from the perspective of traditional DOA estimation methods. Suppose that the signals are independent of each other and that all steering vectors have been normalized, without exception. Then, we can redefine two decoupled covariance matrices for the noiseless observation data as follows:

\[
\begin{align*}
\bar{R}_\theta & \equiv XX^H \\
& = \sum_{k=1}^{K} a_N(\theta_k) a_N^H(\phi_k) s_k(t) s_k^*(t) a_M(\phi_k) a_M^H(\theta_k) \\
\bar{R}_\phi & \equiv X^H X \\
& = \sum_{k=1}^{K} |s_k(t)|^2 a_N(\theta_k) a_N^H(\theta_k) a_M(\theta_k) a_M^H(\theta_k) \\
& = \sum_{k=1}^{K} |s_k(t)|^2 a_M(\phi_k) a_N^H(\theta_k) a_M^H(\phi_k) \\
& = \sum_{k=1}^{K} |s_k(t)|^2 a_M(\phi_k) a_N^H(\theta_k) a_M^H(\phi_k)
\end{align*}
\]

From the proof of Proposition 1, we know that the Toeplitz matrices \(T(u_{\theta})\) and \(T(u_{\phi})\) are exactly equivalent to the two decoupled covariance matrices above except that the coefficients include a squared difference. The whole RDAM process, therefore, can be viewed as the process of recovering \(\bar{R}_\theta\) and \(\bar{R}_\phi\) with the regulators \(\epsilon_\theta\) and \(\epsilon_\phi\). Furthermore, the two atomic weight values in (24) are related to the Capon algorithm; in particular, \(w_\theta^2(\theta)\) and \(w_\phi^2(\varphi)\) happen to be the 1D modified cost function of the Capon algorithm. In other words, we utilize the inverse of the revised power spectrum to weight the atoms, with the aim of penalizing directions without sources. Moreover, note that the atomic weight values produced by the log-det function \(f_1(T(u))\) correspond to the conventional cost function; this comparison, on the other hand, shows that our surrogate \(f_2(T(u))\) performs better.

Remarkably, however, our algorithm demonstrates superiority beyond 2D DOA estimation with respect to conventional subspace methods such as MUSIC and ESPRIT. The 2D MUSIC method not only requires prior information on the number of sources but also incurs a heavy computational load for the 2D spectrum search; by contrast, the proposed algorithm can directly estimate continuous-valued angles through two decompositions, and it merely needs prior knowledge of the noise level (moreover, even this is unnecessary, as shown by a later analysis). Meanwhile, the 2D ESPRIT method is suitable only for an array with the property of so-called geometric invariance, whereas the proposed algorithm can be applied to a sparse rectangular array (SRA) with missing array outputs or only a few elements of a URA and still achieves better estimation performance under certain conditions [12, 15, 19, 35]. In addition, far fewer snapshots than the array size are required for exact estimation, whereas this is not true of the subspace methods.

4. Further Analysis

4.1. Superresolution. The DAM algorithm requires the following angular interval conditions [20] to guarantee that directions can be estimated from the observed data \(X^*\) with probability \(1 - \delta\), where \(\delta\) is the estimation precision.

\[
\begin{align*}
\Delta_\theta & = \min_{i \neq j} |\sin \theta_i - \sin \theta_j| \geq \frac{1.19}{[(N-1)/4]} \\
\Delta_\phi & = \min_{i \neq j} |\sin \varphi_i - \sin \varphi_j| \geq \frac{1.19}{[(M-1)/4]}
\end{align*}
\]

Here, \(\Delta_\theta\) and \(\Delta_\phi\) are the wrap-around distances on the complex unit circle. These conditions mean that the resolutions in terms of \(\theta\) and \(\varphi\) are inversely proportional to the numbers of sensors by row and column, respectively. By contrast, the proposed algorithm (RDAM) has no such restriction on the angular intervals. As the analysis in Section 3 shows, the quadratic inverse of the solution from the previous iteration, denoted by \(W_{\theta,j} = \epsilon_\theta(T(u_{\theta,j}) + \epsilon_\theta I)^{-2}\) and \(W_{\phi,j} = \epsilon_\phi(T(u_{\phi,j}) + \epsilon_\phi I)^{-2}\), is utilized as the weighting factors for the objective function (21) in the current iteration; thus, large weights are applied to small eigenvalues, while small weights are applied to large eigenvalues. Therefore, minimizing (21) is equivalent to penalizing small eigenvalues, making them much smaller while making large ones relatively larger. Thus, it becomes far easier to pick out the eigenvalues corresponding to signals. In other words, the iterative nature of the algorithm allows it to overcome the limits of the minimal angular interval, and, hence, the resolution is greatly improved. In addition, the simulations presented in the next section similarly indicate that, given a proper initialization, our algorithm yields a very high resolution.

4.2. Computational Complexity. Another merit of RDAM is its low complexity. As mentioned in [18], the vector-based atomic norm minimization (vector-ANM) method has a computational complexity as high as \(O((NM)^{3.5} \log(1/\delta))\) due to its PSD constraint of size \((NM + 1) \times (NM + 1)\). However, in the DAM method [20], the two-level Toeplitz matrix is divided into two one-level Toeplitz matrices, thus reducing the computational complexity to \(O((N + M)^{3.5} \log(1/\delta))\).

Fortunately, our proposed RDAM algorithm has the same complexity as DAM in a single iteration; thus the computational efficiency depends merely on the number of iterations. Since such a reweighted iterative algorithm can converge within a few steps [21], the computational cost of RDAM is no greater than \(d \cdot o((N + M)^{3.5} \log(1/\delta))\), where \(d\) denotes the number of iterations; i.e., the complexity is comparable to that of DAM but far below the heavy load of vector-ANM. An empirical run-time comparison is also provided based on the simulations reported below.

4.3. The Noisy Case. In fact, DOA estimation is typically implemented in a noisy environment; therefore, we consider
the following observation model with additive noise $N$, which involves independent entries drawn from $\mathcal{N}(0, \sigma^2)$:

$$Y^o = X + N = \sum_{k=1}^{K} a_N(\theta_k) a_M^H(\varphi_k) + \sigma^2 I$$ \hfill (31)

Then, an iterative denoising optimization problem is obtained as follows:

$$\min_{u_{o,k}, u_{o,l}, X, \eta} \left[ \text{tr} \left( W_{o,j} T(u_o) \right) + \text{tr} \left( W_{p,j} T(u_p) \right) + \lambda \|Y^o - X\|_F^2 \right]$$

$$\text{s.t.} \begin{bmatrix} T(u_p) & X^H \\ X & T(u_o) \end{bmatrix} \succeq 0$$

where $\lambda$ is a regularization factor that keeps a balance between data fitting and a low-rank solution. When $\lambda \geq 2 \|M\|_2$, where $\|\cdot\|_2$ represents the spectral norm, the soft threshold denoising algorithm has the following asymptotic mean square error (MSE) [36]:

$$\frac{1}{(NM)} \|X - X^o\|_F^2 = \mathcal{O} \left( \frac{\sigma^2 K}{NM} \right)$$ \hfill (33)

In this paper, we directly optimize the dual problem [37] of (32) as follows:

$$\min_{A_{j,l}, Z_{j,l}, V_{j,l}} 2 \text{Rtr} \left( Y^o V^H V \right) + 2\eta \|V\|_F$$

$$\text{s.t.} \begin{bmatrix} A & V^H \\ V & Z \end{bmatrix} \succeq 0$$

$$\sum_{n=1}^{N-l} Z_{n,l,n+l} = \sum_{n=1}^{N-l} (W_{o,j} a_{p,n+1})^T, \quad l = 0, \ldots, N - 1$$

$$\sum_{m=1}^{M-k} A_{m,m+k} = \sum_{m=1}^{M-k} (W_{p,j} a_{p,m+1})^T, \quad k = 0, \ldots, M - 1$$

where $\eta \geq \|Y^o - X\|_F$. In practice, we can set an initial value for $\eta$ and then gradually decrease it by a suitable small factor in each iteration; as shown by the simulations reported below, this approach results in better performance.

5. Numerical Simulations

In this section, we carry out numerical simulations conducted to illustrate the performance of the proposed superresolution algorithm (RDAM) in 2D DOA estimation for a rectangular array. In all experiments, the value of $K$ was not given a priori, and a random phase (uniformly distributed between 0 and $2\pi$), associated with a constant magnitude of 1, was set for each received signal. In addition, all SDPs involved were implemented using CVX [38], and the matrix pencil method was adopted for the Vandermonde decomposition of the Toeplitz matrices. In addition, the RDAM algorithm was initialized as discussed in Section 3, and we not only selected suitable initial factors $a$ and $b$, as required, but also iteratively tuned them by decreasing their values by a factor of 0.5 in each subsequent iteration.

First, let us present an intuitive illustration of the higher resolution of RDAM compared with DAM. Suppose that $K = 5$ narrowband signals with directions of (10.1°, 56.12°), (16.52°, 65.08°), (25.01°, 40.25°), (34.07°, 32.25°), and (40.15°, 47.01°) impinge on a rectangular array of size $N \times M = 10 \times 10$. Note that both max$|\Delta\theta|$ and max$|\Delta\varphi|$ are no larger than 10°. Figure 3(a) shows the DAM estimation result, while Figure 3(b) shows the performance of RDAM with initial factors $a = b = 1$; all results were obtained through 100 Monte Carlo experiments. It is apparent that, in the noiseless case, all the angles are estimated correctly and stably by our algorithm, whereas DAM yields unsatisfactory outcomes. Moreover, RDAM is even robust to changes in the signal phases, possibly extending to certain special fields.

Second, let us compare the proposed algorithm with three other methods, namely, DAM, vector-ANM, and vector-based reweighted atomic norm minimization (vector-RAM) [18], to provide a more reliable verification of the superresolution of our algorithm. In particular, let $K = 2$ and $N = M = 10$. We fix the direction of one of the two signals as $(30°, 60°)$, while the other signal is radially separated from the fixed signal by azimuth and elevation spacings that both range from 2° to 20°. Each interval value is allowed to randomly fluctuate within a range of ±0.025°. Let $a = b = 1$, and let the number of iterations of RDAM be four; the same parameter settings are considered for vector-RAM. Figure 4 compares the results of 500 Monte Carlo experiments each for the azimuth $\theta$ and the elevation. It is clear that the proposed algorithm can distinguish spacing angles that differ by 3° in both dimensions, while the other algorithms cannot. Moreover, it achieves better performance than vector-RAM, which is a reweighted algorithm based on stretched vectors, under the same initial conditions. In addition, we list the run-times of these four methods in Table 1, from which we can see that, for both vector-based methods, the computational complexity increases dramatically as the size of the array increases. By contrast, the proposed algorithm and the DAM algorithm both have very low complexity regardless of the array size.

Furthermore, we examine the iterative process of RDAM to investigate the hypothesis that the proposed surrogate function $f_2(T(u))$ enables the number of iterations to be reduced, i.e., more effectively encourages low-rank solutions.
and enables a faster convergence speed. Accordingly, let $K = 3$ and $N = M = 10$, and consider randomly selected signal directions of $(34.35^\circ, 67.03^\circ)$, $(38.09^\circ, 76.04^\circ)$, and $(44.12^\circ, 70.56^\circ)$. Initial factors of 0.01 and 0.1 are selected for the functions $f_1(T(u))$ and $f_2(T(u))$, respectively; under these conditions, the subfunctions have similar approximation properties, as seen from Figure 2. The iterative processes using the two surrogate functions are shown in Figure 5, where the first two subplots in both (a) and (b) present the changes in the eigenvalues for $\theta$ and $\phi$, respectively, and the third subplot shows the estimation performance in each iteration. It is obvious that the previous function $f_1(T(u))$ requires four iterations while the proposed function $f_2(T(u))$ needs only three iterations, which indicates that the proposed function has a faster estimation speed.

Finally, we carry out 200 Monte Carlo experiments considering the case of very closely spaced sources to illustrate the antinoise performance of the proposed algorithm relative to the Cramér-Rao bound (CRB). The array response with i.i.d. additive Gaussian noise satisfying $N(0, \sigma^2)$ was obtained; the signal-to-noise ratio (SNR) was $10 \log \sigma^{-2}$ since the powers of all signals were set to a constant value of 1.
As before, we set \( N = M = 10 \), and the directions of the two sources were set to \((10.10^\circ, 40.01^\circ)\) and \((13.56^\circ, 43.12^\circ)\). Figure 6 shows a graph of the average root mean square error (RMSE) with respect to the SNR with the parameters set to \( a = b = 1 \) and the initial noise level set to \( \sqrt{NM}\sigma^2 \). It is worth noting that the noise level was reduced by a factor of 0.8 in the first weighted iteration for both our RDAM algorithm and the vector-RAM algorithm. Obviously, the performance of the proposed algorithm is close to the CRB at higher SNRs, although its antinoise performance may be worse below SNR=15dB; the main reason for the performance degradation is that, at lower SNRs, the probability of efficient estimation decreases so that a very small amount of deteriorating estimates is obtained, which influences the performance of the curve. However, our algorithm still shows more robust performance than the others in the absence of noise.

6. Conclusion

In this paper, we proposed a superresolution decoupled reweighted atomic norm minimization (RDAM) algorithm for 2D DOA estimation based on the popular theory of continuous compressed sensing (CCS). By directly employing the atomic \( \ell_0 \) norm and converting the optimization problem into a decoupled rank minimization problem, we ensured that the computational complexity is greatly reduced. On the other side, we presented a novel surrogate function and formulated a new concave optimization problem, overcoming the shortcomings of existing methods in terms of resolution and also accelerating the convergence speed. Our algorithm is even suitable for application to sparse rectangular array (SRA), although there is a need for better denoising processing, which remains to be studied in future work.

Appendix

A. Proof of Proposition 2

First, we rewrite the constraint condition as

\[
T^\top (u_\psi) \geq 0, \\
(I_M - T^\top (u_\psi) T (u_\psi)^\top) X^H = 0, \\
T (u_\varphi) - X T (u_\psi)^\top X^H \geq 0
\]
This implies that $X$ lies in the ranges of both $T(u_p)$ and $T(u_0)$, and because $X$ is composed of atoms of the form $\{a_N(\theta_k, 0) a_N^H(\theta_k, 0)\}_k$, we assume the following Vandermonde decomposition for $T(u_p)$ and $T(u_0)$:

$$T(u_0) = \sum_{k} c_{0,k} a_N(\theta_k, 0) a_N^H(\theta_k, 0)$$

$$T(u_p) = \sum_{k} c_{p,k} a_M(\varphi_k, 0) a_M^H(\varphi_k, 0)$$

Let the right-hand side of (25) be denoted by $P(25)$; then, we have

$$P(25) = \min \frac{1}{2} \text{tr} \left( W_\theta \sum_k c_{0,k} a_N(\theta_k) a_N^H(\theta_k) \right) - \frac{1}{2} \text{tr} \left( W_\varphi \sum_k c_{p,k} a_M(\varphi_k) a_M^H(\varphi_k) \right)$$

$$= \min \frac{1}{2} \sum_k c_{0,k} \text{tr} \left( W_\theta a_N(\theta_k) a_N^H(\theta_k) \right)$$

$$+ \frac{1}{2} \sum_k c_{p,k} \text{tr} \left( W_\varphi a_M(\varphi_k) a_M^H(\varphi_k) \right)$$

$$= \min \frac{1}{2} \sum_k c_{0,k} \text{tr} \left( W_\theta a_N(\theta_k) \right)$$

$$+ \frac{1}{2} \sum_k c_{p,k} \text{tr} \left( W_\varphi a_M(\varphi_k) \right)$$

Next, the following conclusion is obtained from the given conditions:

$$P(25) = \min \frac{1}{2} \sum_k c_{0,k} \text{tr} \left( W_\theta a_N(\theta_k) \right)$$

In addition, according to the analysis in Section 3, the following two decompositions hold for the weighted atoms:

$$XX^H = \sum_k |s_{k}^w|^2 w_\theta(\theta_k) w_\varphi^*(\varphi_k) a_N(\theta_k) a_N^H(\theta_k)$$

$$X^HX = \sum_k |s_{k}^w|^2 w_\varphi(\varphi_k) w_\varphi^*(\varphi_k) a_M(\varphi_k) a_M^H(\varphi_k)$$

Thus, we have $c_{0,k} = |s_{k}^w| w_\theta^2(\theta_k)$ and $c_{p,k} = |s_{k}^w| w_\varphi^2(\varphi_k)$. Therefore, $P(25) = \min(1/2) \sum_k |s_{k}^w|^2 + (1/2) \sum_k |s_{k}^w|^2 = \|X\|_{2w}^2$.

B. Derivation of the First-Order Taylor Polynomial (11)

Inspired by the derivations in [39], the partial derivative of the first term of (18) is computed as follows:

$$\partial \left( -\varepsilon_0 \text{tr} \left( (T(u_0) + \varepsilon_0 I)^{-1} \right) \right)$$

$$= -\varepsilon_0 \text{tr} \left( \partial (T(u_0) + \varepsilon_0 I)^{-1} \right) = -\varepsilon_0 \text{tr} \left( - (T(u_0) + \varepsilon_0 I)^{-1} \right)$$

$$= -\varepsilon_0 \text{tr} \left( - (T(u_0) + \varepsilon_0 I)^{-1} \right)$$

$$= -\varepsilon_0 \text{tr} \left( (T(u_0) + \varepsilon_0 I)^{-2} \partial (T(u_0)) \right)$$

Thus, the first-order approximation of this term at the current point $[T(u_{0,i}), T(u_{p,j})]$ can be expressed as follows:

$$-\varepsilon_0 \text{tr} \left( (T(u_0) + \varepsilon_0 I)^{-1} \right)$$

$$= \varepsilon_0 \text{tr} \left( (T(u_{0,i}) + \varepsilon_0 I)^{-1} \right)$$

$$+ \text{tr} \left( \varepsilon_0 (T(u_{0,i}) + \varepsilon_0 I)^{-2} (T(u_0) - T(u_{0,i})) \right)$$

Data Availability

The data used to support the findings of this study are not available because of the following reasons: (1) The data in this paper is based on computer simulation and is generated randomly by itself. (2) This generation method is common in the field of DOA estimation; we consulted the authoritative passages, such as "G. Tang, B. N. Bhaskar, P. Shah, 'Compressed sensing off the grid,' IEEE Trans. Inf. Theory, vol. 59, no. 11, pp. 7465-7490, Nov. 2013," and "Z. Tian, Z. Zhang, Y. Wang, 'Low-complexity optimization for two-dimensional direction-of-arrival estimation via decoupled atomic norm minimization,' in IEEE International Conference on Acoustics, Speech and Signal Processing, 2017."

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported in part by the Natural Science Basic Research Project of Shaanxi Province under Grant 2018JQ6046 and in part by the Fundamental Research Funds for the Central Universities of China under Grant XJS18033.
References


