Distributed Sparse Spectral Estimation: From Multilevel Reconstruction Viewpoint

Haoyue Zhang
College of Electronic Countermeasure
National University of Defense Technology
Hefei, China
zhanghaoyue23@163.com

Junpeng Shi
College of Electronic Countermeasure
National University of Defense Technology
Hefei, China
shijunpeng20@nudt.edu.cn

Shuyun Shi
College of Electronic Countermeasure
National University of Defense Technology
Hefei, China
13478699873@163.com

Yi Ding
College of Electronic Countermeasure
National University of Defense Technology
Hefei, China
dingyi200005@163.com

***Abstract*—This paper addresses distributed line spectral estimation problem arising in multi-node sampling systems. While recent studies have explored the distributed compressed sensing and shown its advantages in terms of data compression, few works consider the intrinsic signal correlation to achieve joint sparse spectral estimation. We thus propose the distributed covariance fitting (DCF) method by employing the joint sparsity and cross correlation to reconstruct the multilevel structured matrix and extract unknown parameters. We formulate a semidefinite program (SDP) via the covariance fitting criteria where the structured feature of the second-order statistics for signal ensemble is exploited. The connection between the DCF method and classical atomic norm minimization method illustrates the effective application of joint sparsity and redundant information of the signal ensemble. Extensive numerical results are provided that corroborate our analysis and exhibit the superior performance of proposed method.**

***Index Terms*—Distributed line spectral estimation, joint sparsity, inter-signal correlation, multilevel reconstruction.**

I. Introduction

LINE spectral estimation is a classic problem in statistical signal processing [1]. It generally aims at estimating spectral parameters from noisy sinusoidal signals sampled from a single node. In this paper, we consider the distributed line spectral estimation problem which is devoted to estimating multi-dimensional spectral parameters from a set of noisy signals sampled from multiple nodes and sharing the same complex amplitude information. This problem has extensive applications in array processing [2] and radar [3], [4].

Subspace-based methods [5], [6] and their extensions [7], [8] are widely used for line spectral estimation, even in distributed systems to cope with colored noise [4], [9]. However, they do not essentially extend the array aperture, and therefore are not able to improve the estimation accuracy under typical operative conditions. Besides, subspace-based methods may degrade or even fail under limited channels or low signal to noise ratio (SNR) conditions in practice [10]. With the development of compressed sensing, the gridless sparse methods have gained widespread popularity in line spectral estimation [11]–[13] due to their ability to circumvent the grid mismatch problem caused by traditional sparse recovery framework [14]. The typical sparse approaches include deterministic methods, e.g., atomic norm minimazation (ANM) and its extensions [15], [16], and covariance fitting methods [17], [18], which are all capable of implementing the spectral estimation directly in the continous frequency domain. Meanwhile, they maintain high parameter resolution under aforementioned challenging conditions when subspace-based methods degrade severely.

Recently, distributed compressed sensing (DCS) provides a systemic joint sparse framework for the signal ensemble [19], employing the joint sparsity and inter-signal correlation to achieve signal recovery with smaller measurements. However, DCS is devoted to randomly generating the sensing matrices to compress the data, thus not suitable for line spectral estimation. Then, how to combine the DCS framework with distributed line spectral estimation to fully exploit the joint sparsity and redundancy information is a significant challenge.

In this paper, we propose a distributed covariance fitting (DCF) method for distributed line spectral estimation and formulate the corresponding optimization problem by reconstructing distributed covariance matrix via its connection to multilevel Toeplitz (MLT) structure. The proposed method effectively exploits the joint sparsity and redundant information among the signal ensemble to achieve multilevel reconstruction, offering a way to improve the estimation performance. Numercial results are provided to demonstrate the superior performance of the DCF method.

Notations: Boldface letters are reserved for vectors and matrices. $\mathrm{tr}(\boldsymbol{X})$ means the trace of $\boldsymbol{X}$, respectively. The i-th row (j-th column) of $\boldsymbol{X}$ is $\boldsymbol{X}(i,:)$ ($\boldsymbol{X}(:,j)$).

The research of the project was supported by the National Natural Science Foundation of China under Grant 62071476. (*Corresponding author: Junpeng Shi*)

II. SIGNAL MODEL

We consider a array direction finding problem with K narrowband sources impings on N uniform linear arrays configured orthogonally at the same location for application, then the data at the t-th snapshot sampled from the n-th subarray can be represented as

$$\boldsymbol{Y}_n(t)=\sum_{k=1}^{K}\boldsymbol{a}_n(f_{nk})s_{nk}(t)+\boldsymbol{e}_n(t), n=1,\dots,N, \quad (1)$$

where $i=\sqrt{-1}$, $\boldsymbol{a}_n(f)=\left[1,e^{i2\pi f},\dots,e^{i2\pi(M_n-1)f}\right]^T$, M_n denotes the number of samples at the n-th node and $t\in[L]:=\{1,2,\dots,L\}$; $s_k(t)\in\mathbb{C}$ and $f_k\in[0,1)$ are the amplitude and frequency of the k-th component respectively and $\boldsymbol{e}_n(t)$ is the measurement noise. Specially, two-dimensional (2D) direction of arrival (DOA) estimation for L-shaped array can be considered a two-nodes case. Then, with the shared source signal $\boldsymbol{S}$, all the data can be gathered as follows

$$\boldsymbol{Y}=\begin{bmatrix}\boldsymbol{A}_1(\boldsymbol{f}_1)\\ \boldsymbol{A}_2(\boldsymbol{f}_2)\\ \vdots\\ \boldsymbol{A}_N(\boldsymbol{f}_N)\end{bmatrix}\boldsymbol{S}+\begin{bmatrix}\boldsymbol{E}_1\\ \boldsymbol{E}_2\\ \vdots\\ \boldsymbol{E}_N\end{bmatrix}\in\mathbb{C}^{M\times L}, \quad (2)$$

where $M=\sum_{n=1}^{N}M_n$ denotes the number of all the samples, $\boldsymbol{f}_n=\begin{bmatrix}f_{n1} & \dots & f_{nK}\end{bmatrix}^T$ are frequencies at the n-th node, and $\boldsymbol{A}_n(\boldsymbol{f}_n)=\begin{bmatrix}\boldsymbol{a}_n(f_{n1}) & \dots & \boldsymbol{a}_n(f_{nK})\end{bmatrix}$ denotes the corresponding manifold matrix. For notational simplicity, we will write $\boldsymbol{A}_n(\boldsymbol{f}_n)$, $\boldsymbol{a}_{nk}$ as $\boldsymbol{A}_n$ and $\boldsymbol{a}_n(f_{nk})$ hereafter.

Assume that $\boldsymbol{S}$ and $\boldsymbol{E}$ are both independent random variables, and they are uncorrelated with each other. Besides, $\boldsymbol{N}_n$ is additive zero means Gaussian white noise and every channel of $\boldsymbol{S}$ obeys zero means complex Gaussian distribution, i.e.,

$$\begin{aligned}\mathbb{E}(\boldsymbol{S}(:,m)\boldsymbol{S}(:,n)^H)&=\begin{cases}\boldsymbol{P}, & m=n\\ \boldsymbol{0}, & m\neq n\end{cases},\\ \mathbb{E}(\boldsymbol{E}(:,m)\boldsymbol{E}(:,n)^H)&=\begin{cases}\operatorname{diag}(\boldsymbol{\sigma}), & m=n\\ \boldsymbol{0}, & m\neq n\end{cases},\end{aligned} \quad (3)$$

where $\boldsymbol{P}=\operatorname{diag}(p_1,\cdots,p_K)$, $p_k>0$ and $\boldsymbol{\sigma}\in\mathbb{C}^M$ always hold in practice. Then the distributed covariance matrix of $\boldsymbol{Y}$ can be represented as

$$\begin{aligned}\hat{\boldsymbol{R}}&=\begin{bmatrix}\boldsymbol{A}_1\\ \vdots\\ \boldsymbol{A}_N\end{bmatrix}\boldsymbol{P}\begin{bmatrix}\boldsymbol{A}_1\\ \vdots\\ \boldsymbol{A}_N\end{bmatrix}^H+\operatorname{diag}(\boldsymbol{\sigma})\\ &=\begin{bmatrix}\boldsymbol{R}_{11} & \cdots & \boldsymbol{R}_{1N}\\ \vdots & \ddots & \vdots\\ \boldsymbol{R}_{N1} & \cdots & \boldsymbol{R}_{NN}\end{bmatrix}+\operatorname{diag}(\boldsymbol{\sigma})=\boldsymbol{R}+\operatorname{diag}(\boldsymbol{\sigma}),\end{aligned} \quad (4)$$

where $\boldsymbol{R}$ denotes the noise-free covariance matrix, $\boldsymbol{R}_{mn}\in\mathbb{C}^{M_m\times M_n}$ is the (m,n)-th block matrix of $\boldsymbol{R}$, and the sample covariance matrix can be calculated by

$$\tilde{\boldsymbol{R}}=\frac{1}{L}\boldsymbol{Y}\boldsymbol{Y}^H. \quad (5)$$

III. DISTRIBUTED COVARIANCE FITTING

It is intuitive to derive the similar covariance fitting criteria for distributed line spectral estimation problem as single node case by generalized least squares method [20], [21] as

$$h_1=\|\hat{\boldsymbol{R}}^{-\frac{1}{2}}(\tilde{\boldsymbol{R}}-\hat{\boldsymbol{R}})\tilde{\boldsymbol{R}}^{-\frac{1}{2}}\|_{\mathrm{F}}^2, \quad (6)$$

when $L>M$ which means that the inverse of $\tilde{\boldsymbol{R}}$ exists. The key point of this method is to characterize the structure of distributed covariance matrix. Then we will study its structure under typical scenario as $K<\min\{M_n\}_{n=1}^N$.

The ideal distributed covariance matrix $\boldsymbol{R}$ is no longer a Toeplitz matrix as that in single node case, but it still has excellent structural property and can be reconstructed from the MLT structure [22]. According to (4), all the block matrices of $\boldsymbol{R}$ can be represented as

$$\boldsymbol{R}_{mn}=\boldsymbol{A}_m\boldsymbol{P}\boldsymbol{A}_n^H, m,n=1,\dots,N. \quad (7)$$

Then we can derive the following proposition to accurately characterize the structure of the distributed covariance matrix.

Proposition 1: Consider a positive semidefinite (PSD) MLT matrix $\boldsymbol{T}_N$ which can be uniquely decomposed with the sufficient condition that $\operatorname{rank}(\boldsymbol{T}_N)=r<\min\{M_n\}_{n=1}^N$ as [22, Theorem 1]

$$\boldsymbol{T}_N=(\boldsymbol{A}_N\circ\cdots\circ\boldsymbol{A}_1)\boldsymbol{P}(\boldsymbol{A}_N\circ\cdots\circ\boldsymbol{A}_1)^H, \quad (8)$$

where $\circ$ denotes the Khatri-Rao product, then every block matrix of ideal distributed covariance matrix $\boldsymbol{R}$ which satisfies $\operatorname{rank}(\boldsymbol{R})=r$ can be constructed by $\boldsymbol{T}_N$ as

$$\boldsymbol{R}_{mn}=\boldsymbol{A}_m\boldsymbol{P}\boldsymbol{A}_n^H=\boldsymbol{T}_N(1{:}N_{m-1}{:}N_m,1{:}N_{n-1}{:}N_n), \quad (9)$$

where $N_q=\begin{cases}\prod_{p=1}^q M_p, & q\geq 1\\ 1, & q=0\end{cases}$.

Proof: For ease of notation, we will write the MLT matrix $\boldsymbol{T}_N$ as $\boldsymbol{T}$ in the proof. Without loss of generality, assume that $n\geq m$ in $\boldsymbol{R}_{mn}$ and let $\boldsymbol{T}_q\in\mathbb{C}^{N_q\times N_q}$ denotes a submatrix of $\boldsymbol{T}$ as follows

$$\begin{aligned}\boldsymbol{T}_q&=\boldsymbol{T}(1:N_q,1:N_q)\\ &=\sum_{k=1}^{r}p_k(\boldsymbol{a}_{qk}\otimes\cdots\otimes\boldsymbol{a}_{1k})(\boldsymbol{a}_{qk}\otimes\cdots\otimes\boldsymbol{a}_{1k})^H\\ &=\sum_{k=1}^{r}p_k(\boldsymbol{a}_{qk}\boldsymbol{a}_{qk}^H\otimes\cdots\otimes\boldsymbol{a}_{1k}\boldsymbol{a}_{1k}^H),\end{aligned} \quad (10)$$

where $\otimes$ denotes the Kronecker product and $\operatorname{rank}(\boldsymbol{T})=r<\min\{M_n\}_{n=1}^N$, then we can define $\boldsymbol{Q}_{qk}$ as

$$\boldsymbol{Q}_{qk}=(\boldsymbol{a}_{qk}\boldsymbol{a}_{qk}^H\otimes\cdots\otimes\boldsymbol{a}_{1k}\boldsymbol{a}_{1k}^H)\in\mathbb{C}^{N_q\times N_q}, q\leq N, \quad (11)$$

where $\boldsymbol{a}_{jk}=[1,\phi_{jk},\cdots,\phi_{jk}^{(M_j-1)}], \phi_{jk}=e^{i2\pi f_{jk}}, j=1,2,\dots,q$. Then $\boldsymbol{Q}_{nk}$ can be calculated recursively by

$$\begin{aligned}\boldsymbol{Q}_{nk}&=(\boldsymbol{a}_{nk}\boldsymbol{a}_{nk}^H\otimes\boldsymbol{Q}_{(n-1)k})\\ &=\begin{bmatrix}1\boldsymbol{Q}_{(n-1)k} & \phi_{nk}^{-1}\boldsymbol{Q}_{(n-1)k} & \cdots & \phi_{nk}^{1-n}\boldsymbol{Q}_{(n-1)k}\\ \phi_{nk}\boldsymbol{Q}_{(n-1)k} & 1\boldsymbol{Q}_{(n-1)k} & \cdots & \phi_{nk}^{2-n}\boldsymbol{Q}_{(n-1)k}\\ \vdots & \vdots & \ddots & \vdots\\ \phi_{nk}^{n-1}\boldsymbol{Q}_{(n-1)k} & \phi_{nk}^{n-2}\boldsymbol{Q}_{(n-1)k} & \cdots & 1\boldsymbol{Q}_{(n-1)k}\end{bmatrix}.\end{aligned} \quad (12)$$

It is clear that the first N_{n-1} rows of $\boldsymbol{Q}_{nk}$ are equal to $\boldsymbol{a}_{nk}^H \otimes \boldsymbol{Q}_{(n-1)k} \in \mathbb{C}^{N_{n-1}\times n}$. Therefore, we can get $\boldsymbol{a}_{nk}^H$ by taking interval values from the first row of $\boldsymbol{Q}_{nk}$ as

$$\boldsymbol{a}_{nk}^H = \boldsymbol{Q}_{nk}(1, 1 : N_{n-1} : N_n). \tag{13}$$

Besides, under the assumption that $n \geq m$, $\boldsymbol{Q}_{mk} = \boldsymbol{Q}_{nk}(1 : N_m, 1 : N_m)$, then it is shown that

$$\boldsymbol{a}_{mk} = \boldsymbol{Q}_{nk}(1 : N_{m-1} : N_m, 1). \tag{14}$$

According to the results above, it follows that

$$\begin{aligned}\boldsymbol{R}_{mn} &= \sum_{k=1}^{r} p_k \boldsymbol{a}_{mk}\boldsymbol{a}_{nk}^H \\ &= \sum_{k=1}^{r} p_k \boldsymbol{Q}_{nk}(1 : N_{m-1} : N_m, 1)\boldsymbol{Q}_{nk}(1, 1 : N_{n-1} : N_n) \\ &= \sum_{k=1}^{r} p_k \boldsymbol{Q}_{Nk}(1 : N_{m-1} : N_m, 1)\boldsymbol{Q}_{Nk}(1, 1 : N_{n-1} : N_n).\end{aligned} \tag{15}$$

Let $\boldsymbol{W}_k = \boldsymbol{a}_{Nk} \otimes \cdots \otimes \boldsymbol{a}_{1k}$, then it holds that

$$\boldsymbol{Q}_{Nk}(p, 1) = \boldsymbol{W}_k(p), \quad \boldsymbol{Q}_{Nk}(1, q) = \boldsymbol{W}_k(q)^*. \tag{16}$$

Then for elements of $\boldsymbol{T}$, it always holds that

$$\begin{aligned}\boldsymbol{T}(p, q) &= \sum_{k=1}^{r} p_k \boldsymbol{W}_k(p)\boldsymbol{W}_k(q)^* \\ &= \sum_{k=1}^{r} p_k \boldsymbol{Q}_{Nk}(p, 1)\boldsymbol{Q}_{Nk}(1, q),\end{aligned} \tag{17}$$

and it follows that

$$\boldsymbol{R}_{mn} = \boldsymbol{T}(1 : N_{m-1} : N_m, 1 : N_{n-1} : N_n). \tag{18}$$

Moreover, it always holds that

$$\boldsymbol{T}_{mn} = \boldsymbol{T}_{nm}^H. \tag{19}$$

Therefore, Proposition 1 is proved. ∎

Based on the proposition and the covariance fitting criteria defined in (6), we can naturally formulate the minimization of h_1 as the SDP problem from multilevel reconstruction viewpoint under general $L > M$ case as follows:

$$\begin{aligned}&\min_{\boldsymbol{X}, \boldsymbol{T}_N, \boldsymbol{\sigma}\succeq \boldsymbol{0}} \operatorname{tr}(\boldsymbol{X}) + \operatorname{tr}(\tilde{\boldsymbol{R}}^{-1}\hat{\boldsymbol{R}}) \\ &\text{s.t.} \begin{bmatrix} \boldsymbol{X} & \tilde{\boldsymbol{R}}^{\frac{1}{2}} \\ \tilde{\boldsymbol{R}}^{\frac{1}{2}} & \hat{\boldsymbol{R}} \end{bmatrix} \succeq \boldsymbol{0}, \boldsymbol{T}_N \succeq \boldsymbol{0}, \operatorname{rank}(\boldsymbol{T}_N) \leq \min\{M_n\}_{n=1}^N.\end{aligned} \tag{20}$$

where $\hat{\boldsymbol{R}}$ is constructed from (4) and (9), and the rank constraint is a sufficient condition which is tight to ensure the unique Vandermonde decomposition of the MLT matrix [22]. However, the rank constraint is non-convex and may limit the maximum number of detectable components.

Inspired by the SPA method in the single node case [17], we relax the rank constraint and reformulate the problem as

$$\begin{aligned}&\min_{\boldsymbol{X}, \boldsymbol{T}_N, \boldsymbol{\sigma}\succeq \boldsymbol{0}} \operatorname{tr}(\boldsymbol{X}) + \operatorname{tr}(\tilde{\boldsymbol{R}}^{-1}\hat{\boldsymbol{R}}) \\ &\text{s.t.} \begin{bmatrix} \boldsymbol{X} & \tilde{\boldsymbol{R}}^{\frac{1}{2}} \\ \tilde{\boldsymbol{R}}^{\frac{1}{2}} & \hat{\boldsymbol{R}} \end{bmatrix} \succeq \boldsymbol{0}, \boldsymbol{T}_N \succeq \boldsymbol{0}.\end{aligned} \tag{21}$$

Upon solving the problem, the optimal PSD MLT matrix can be represented as

$$\boldsymbol{T}_N^* = (\boldsymbol{A}_N \circ \cdots \boldsymbol{A}_1)\boldsymbol{P}(\boldsymbol{A}_N \circ \cdots \boldsymbol{A}_1)^H + \delta\boldsymbol{I}, \tag{22}$$

where $\delta\boldsymbol{I}$ is generated due to the redundant variable on the principal diagonal of $\hat{\boldsymbol{R}}$, and it follows that $\operatorname{rank}(\boldsymbol{T}_N^*) \geq K$. Then the solution of (20) is one special realization of (22) as $\delta = 0$, and we can implement the postprocessing step to maintain only the signal subspace by calculate K-order truncated eigendecomposition of $\boldsymbol{T}_N^*$. Although the truncated eigendecomposition may destroy the MLT structure of $\boldsymbol{T}_N^*$, making the Mapp algorithm [22] unusable, the classical multidimensional ESPRIT (MD-ESPRIT) algorithm can be used directly for frequency extraction and pairing [23].

Till now, we have finished the distributed line spectral estimation with the proposed DCF algorithm and the detailed steps are summarized as Algorithm 1.

Algorithm 1 Distributed Covariance Fitting (DCF)

Require:
observed data $\boldsymbol{Y}$ and model order K.
Ensure:
Estimated frequency groups $\{\boldsymbol{f}_k\}_{k=1}^K$, $\boldsymbol{f}_k = (f_{1k}, \ldots, f_{Nk})$.
1: Calculate the sample covariance matrix $\tilde{\boldsymbol{R}}$ as (5);
2: Estimate $\boldsymbol{T}_N^*$ by solving (21) with SDP3 solver;
3: Postprocess $\boldsymbol{T}_N^*$ with K-order truncated eigendecomposition to obtain $\boldsymbol{T}_N^K$;
4: Extract and pair the frequency groups using MD-ESPRIT algorithm from $\boldsymbol{T}_N^K$;
5: **return** $\{\boldsymbol{f}_k\}_{k=1}^K$.

Similar to the single node case, the DCF method also has close connection to the ANM framework.

Lemma 1: The SDP problem in (21) is equivalent to the following reweighted ANM (RAM) formulation as

$$\min_{\boldsymbol{Z}} \sqrt{M}\|\boldsymbol{Z}\|_{\mathcal{A}^w} + \sum_{m=1}^{M} \sqrt{(\tilde{\boldsymbol{R}}^{-1})_{ii}}\|(\tilde{\boldsymbol{R}}^{\frac{1}{2}} - \boldsymbol{Z})_i\|_2, \tag{23}$$

where $\|\boldsymbol{Z}\|_{\mathcal{A}^w}$ denotes the weighted atomic norm [24] defined in general case as

$$\|\boldsymbol{Z}\|_{\mathcal{A}^w} := \inf_{\boldsymbol{f}_k, \boldsymbol{s}_k} \{\sum_{k=1}^{K} \frac{\|\boldsymbol{s}_k\|_2}{w_k} : \boldsymbol{Z} = \sum_{k=1}^{K} \boldsymbol{a}(\boldsymbol{f}_k)\boldsymbol{s}_k\}, \tag{24}$$

where

$$\boldsymbol{a}(\boldsymbol{f}_k) = \begin{bmatrix}\boldsymbol{a}_{1k}^T & \boldsymbol{a}_{2k}^T & \cdots & \boldsymbol{a}_{Nk}^T\end{bmatrix}^T, \tag{25}$$

$$w_k = [\frac{1}{M}\boldsymbol{a}(\boldsymbol{f}_k)^H \boldsymbol{W}\boldsymbol{a}(\boldsymbol{f}_k)]^{-\frac{1}{2}}, \boldsymbol{W} = \tilde{\boldsymbol{R}}^{-1}. \tag{26}$$

Proof: Suppose that $\boldsymbol{R}$ can be precisely constructed as (9), then it can be cast as the following SDP problem:

$$\begin{aligned}\|\boldsymbol{Z}\|_{\mathcal{A}^w} = &\min_{\boldsymbol{X}, \boldsymbol{R}} \frac{1}{2\sqrt{M}}[\operatorname{tr}(\boldsymbol{X}) + \operatorname{tr}(\boldsymbol{W}\boldsymbol{R})], \\ &\text{s.t.} \quad \begin{bmatrix} \boldsymbol{X} & \boldsymbol{Z}^H \\ \boldsymbol{Z} & \boldsymbol{R} \end{bmatrix} \succeq \boldsymbol{0},\end{aligned} \tag{27}$$

which can be proved similar as [24, Theorem 3].

According to [13, lemma4], the problem in (21) can be transformed as follows:

$$\begin{aligned}
&\min_{\boldsymbol{R},\boldsymbol{\sigma}\succeq\mathbf{0}} \operatorname{tr}(\frac{1}{L}\boldsymbol{Y}^H\hat{\boldsymbol{R}}^{-1}\boldsymbol{Y}) + \operatorname{tr}(\tilde{\boldsymbol{R}}^{-1}\hat{\boldsymbol{R}}) \\
\Leftrightarrow &\min_{\boldsymbol{Z},\boldsymbol{R},\boldsymbol{\sigma}} \operatorname{tr}(\boldsymbol{Z}^H\boldsymbol{R}^{-1}\boldsymbol{Z}) + \sum_{i=1}^{M}\frac{1}{\sigma_i}\|(\frac{1}{\sqrt{L}}\boldsymbol{Y}-\boldsymbol{Z})_i\|_2^2 \\
&+ \operatorname{tr}(\tilde{\boldsymbol{R}}^{-1}\boldsymbol{R}) + \sum_{i=1}^{M}\sigma_i(\tilde{\boldsymbol{R}}^{-1})_{ii} \\
\Leftrightarrow &\min_{\boldsymbol{Z}} \sqrt{M}\|\boldsymbol{Z}\|_{\mathcal{A}^w} + \sum_{i=1}^{M}\sqrt{(\tilde{\boldsymbol{R}}^{-1})_{ii}}\|(\frac{1}{\sqrt{L}}\boldsymbol{Y}-\boldsymbol{Z})_i\|_2,
\end{aligned} \tag{28}$$

where $\boldsymbol{W} = \tilde{\boldsymbol{R}}^{-1}$ and $(1/\sqrt{L})\boldsymbol{Y}$ can be substituted by $\tilde{\boldsymbol{R}}^{\frac{1}{2}}$. The lemma is proved. ■

According to (23), the joint sparsity is exhibited in the former item, and the coherence property is characterized by the latter item which calculates the whole reconstruction error of the distributed covariance matrix.

IV. Numercial Results

In this section, we illustrate the performance of the proposed DCF method by comparing it with ESPRIT and SPA [17] algorithm directly acting on $\tilde{\boldsymbol{R}}$ via numerical simulations. Besides, the Cramér-Rao bound (CRB) of algorithms are calculated as the benchmarks. As the DCF method can utilize the correlation information embedded in the distributed covariance matrix, it posseses lower CRB which is represented as CRB+ while that of algorithms performing independently at local node is denoted as CRB.

In the experiments, we randomly generate the complex amplitudes $\{s_k(t)\}$ independently and identically from a standard complex normal distribution for all nodes, and form different frequency gruops $\boldsymbol{f}_k = \{f_{nk}\}_{n=1}^{N}$ for different components by angles as $f_{nk} = \cos(\theta_{nk})/2, \theta_{nk} \in [0, 180°)$ for application in array direction finding. The noise is complex Gaussian white noise with zero mean and the same variance at samples of the same node. The SNR of the n-th node is defined as $10\log_{10}(\|\boldsymbol{Y}_n - \boldsymbol{E}_n\|_F^2/\|\boldsymbol{E}_n\|_F^2)$. Besides, we will mainly focus on the two nodes case for brevity, and the similar results can be obtained for more nodes.

To investigate the performance improvement of the DCF method for frequency identification, we employ samples with $L = 1\times10^6$ in the high SNR regime as $\text{SNR}_1 = \text{SNR}_2 = 30$dB to obtain a relatively ideal sample covariance matrix. Fig. 1 shows the parameter identification performance of proposed method in the case that the sample size is set as $M_1 = M_2 = 5$. It is shown that the proposed method can identify approximately 14 targets exceeding the sensor number of any node in the system, which outperforms the conventional subspace methods and the general algorithms depending on the single node data.

We study the RMSE performance of the proposed method versus the SNR. We fix $L = 200$, $M_1 = M_2 = 6$, $K = 2$ and consider the case that the SNR of two nodes varies simultaneously in $\{-15, -10, \ldots, 10\}$dB. Without loss of generality, $K = 2$ uncorrelated components received by the nodes with $\boldsymbol{\theta}_1 = \{60, 90\}$, $\boldsymbol{\theta}_2 = \{50, 80\}$. $J = 200$ Monte Carlo trials are carried out and the root mean square error (RMSE) of the system can be calculated as

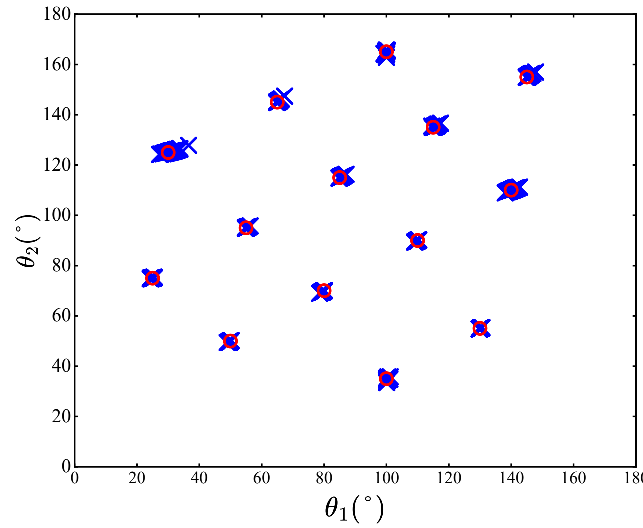
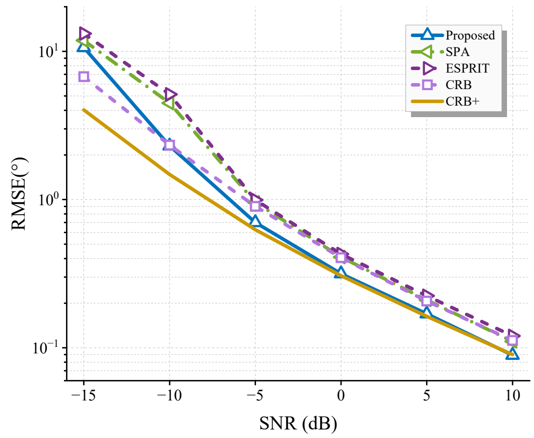


Fig. 1: Estimates of the DCF method with 100 trials, where the sample size is set as $M_1 = M_2 = 5$ (red circles account for the true values and blue crosses stand for the estimates).

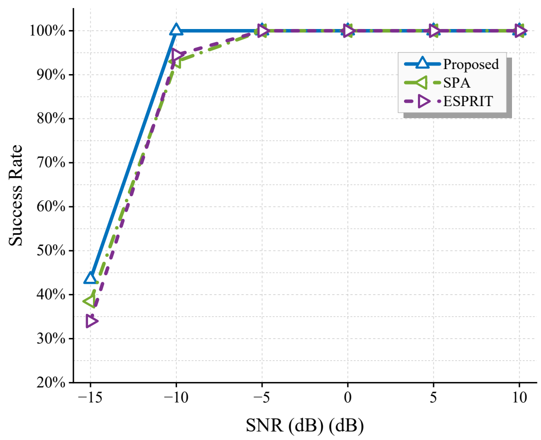
$$\text{RMSE} = \sqrt{\frac{1}{NJK}\sum_{n=1}^{N}\sum_{i=1}^{J}\sum_{k=1}^{K}(\hat{\theta}_{nk}^{(j)} - \theta_{nk})^2}, \tag{29}$$

where $\hat{\theta}_{nk}^{(j)}$ denotes the corresponding estimate of the k-th component for the j-th trial at the n-th node. As few outliers may destroy the reliablity of the experiment, we set a threshold as 30° to measure the success of single experiment in Monte-Carlo runs, which means that if there exists any component with $\left|\hat{\theta}_{nk}^{(j)} - \theta_{nk}\right| > 30°$, the j-th experiment will be considerd a failure, and it will not be included in the calculation of RMSE.

It can be seen that all the algorithms have good performance in the high SNR regime, while the DCF method performs better for it can effectively utilize the redundant imformation embeded in the inter-correlation matrix and so possesses lower CRB. In the moderate/low SNR regime when the measurement noise may cause the leakage of noise to signal subspace and destroy the orthogonality between the signal and noise subspace, the performance of algorithms degrades. However, the DCF method exhibits stronger robustness to noise as it outperforms other methods in RMSE and success rate at -10dB. In the extremely low SNR regime as -15dB, all the algorithms may fail but the proposed method still performs slightly better.



(a) RMSE



(b) Success Rate

Fig. 2: RMSE of estimate results of the whole system. Some settings: $M_1 = M_2 = 6$, $K = 2$, $L = 200$, $\boldsymbol{\theta}_1 = \{60, 90\}$ and $\boldsymbol{\theta}_2 = \{50, 80\}$ (CRB and CRB+ denote benchmarks for independent estimation at local node and multi-node jointly estimation, respectively.)

V. Conclusion

In this paper, we construct a distributed line spectral estimation model in array processing scenario, and propose a sparse method, termed DCF, via the connection between the distributed model and the MLT structure to exploit the joint sparsity and redundant information embedded in the cross-corrlation matrix. Numerical results demonstrate the superior performance of proposed method in terms of frequency identification and estimation accuracy in the low SNR regime.

References

[1] P. Stoica and R. L. Moses, *Spectral Analysis of Signals.* Upper Saddle River, NJ, USA: Pearson/Prentice Hall, 2005.

[2] C. Steffens and M. Pesavento, "Block- and rank-sparse recovery for direction finding in partly calibrated arrays," *IEEE Transactions on Signal Processing*, vol. 66, no. 2, pp. 384–399, Jan. 2018.

[3] E. Fishler, A. Haimovich, R. Blum, L. Cimini, D. Chizhik, and R. Valenzuela, "Spatial Diversity in Radars—Models and Detection Performance," *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 823–838, Mar. 2006.

[4] H. Jiang, J.-K. Zhang, and K. M. Wong, "Joint DOD and DOA Estimation for Bistatic MIMO Radar in Unknown Correlated Noise," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 11, pp. 5113–5125, Nov. 2015.

[5] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, Mar. 1986.

[6] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 7, pp. 984–995, Jul. 1989.

[7] P. Strobach, "Fast recursive subspace adaptive ESPRIT algorithms," *IEEE Transactions on Signal Processing*, vol. 46, no. 9, pp. 2413–2430, Sep. 1998.

[8] W. Liao, "MUSIC for multidimensional spectral estimation: Stability and super-resolution," *IEEE Transactions on Signal Processing*, vol. 63, no. 23, pp. 6395–6406, Dec. 2015.

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

[10] D. Malioutov, M. Cetin, and A. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 3010–3022, Aug. 2005.

[11] B. N. Bhaskar, G. Tang, and B. Recht, "Atomic Norm Denoising With Applications to Line Spectral Estimation," *IEEE Transactions on Signal Processing*, vol. 61, no. 23, pp. 5987–5999, Dec. 2013.

[12] F. Andersson, M. Carlsson, J.-Y. Tourneret, and H. Wendt, "A new frequency estimation method for equally and unequally spaced data," *IEEE Transactions on Signal Processing*, vol. 62, no. 21, pp. 5761–5774, Nov. 2014.

[13] Z. Yang and L. Xie, "On Gridless Sparse Methods for Line Spectral Estimation From Complete and Incomplete Data," *IEEE Transactions on Signal Processing*, vol. 63, no. 12, pp. 3139–3153, Jun. 2015.

[14] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht, "Compressed Sensing Off the Grid," *IEEE Transactions on Information Theory*, vol. 59, no. 11, pp. 7465–7490, Nov. 2013.

[15] Y. Chi and Y. Chen, "Compressive Two-Dimensional Harmonic Retrieval via Atomic Norm Minimization," *IEEE Transactions on Signal Processing*, vol. 63, no. 4, pp. 1030–1042, Feb. 2015.

[16] J. Fang, J. Li, Y. Shen, H. Li, and S. Li, "Super-Resolution Compressed Sensing: An Iterative Reweighted Algorithm for Joint Parameter Learning and Sparse Signal Recovery," *IEEE Signal Processing Letters*, vol. 21, no. 6, pp. 761–765, Jun. 2014.

[17] Z. Yang, L. Xie, and C. Zhang, "A Discretization-Free Sparse and Parametric Approach for Linear Array Signal Processing," *IEEE Transactions on Signal Processing*, vol. 62, no. 19, pp. 4959–4973, Oct. 2014.

[18] S. Sedighi, B. S. M. R. Rao, and B. Ottersten, "An asymptotically efficient weighted least squares estimator for co-array-based DoA estimation," *IEEE Transactions on Signal Processing*, vol. 68, pp. 589–604, 2020.

[19] D. Baron, M. F. Duarte, M. B. Wakin, S. Sarvotham, and R. G. Baraniuk, "Distributed Compressive Sensing," Jan. 2009.

[20] B. Ottersten, P. Stoica, and R. Roy, "Covariance Matching Estimation Techniques for Array Signal Processing Applications," *Digital Signal Processing*, vol. 8, no. 3, pp. 185–210, 1998.

[21] P. Stoica, P. Babu, and J. Li, "SPICE: A Sparse Covariance-Based Estimation Method for Array Processing," *IEEE Transactions on Signal Processing*, vol. 59, no. 2, pp. 629–638, Feb. 2011.

[22] Z. Yang, L. Xie, and P. Stoica, "Vandermonde Decomposition of Multilevel Toeplitz Matrices With Application to Multidimensional Super-Resolution," *IEEE Transactions on Information Theory*, vol. 62, no. 6, pp. 3685–3701, 2016.

[23] C. Jinli, G. Hong, and S. Weimin, "Angle estimation using ESPRIT without pairing in MIMO radar," *Electronics Letters*, vol. 44, no. 24, p. 1422, 2008.

[24] Z. Yang and L. Xie, "Enhancing Sparsity and Resolution via Reweighted Atomic Norm Minimization," *IEEE Transactions on Signal Processing*, vol. 64, no. 4, pp. 995–1006, 2016.