Signal Processing 90 (2010) 207–216

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

A fast algorithm for joint two-dimensional direction of arrival and frequency estimation via hierarchical space–time decomposition ☆

Chun-Hung Lin [a], Wen-Hsien Fang [a,*], Jen-Der Lin [b], Kuo-Hsiung Wu [a]

[a] Department of Electronic Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan, ROC
[b] Department of Electrical Engineering, Nan Jeon Institute of Technology, Tainan, Taiwan, ROC

ARTICLE INFO

Article history:
Received 3 December 2007
Received in revised form
12 May 2009
Accepted 9 June 2009
Available online 17 June 2009

Keywords:
2-D DOA estimation
Carrier frequency estimation
Joint parameter estimation
Unitary ESPRIT
Fast algorithm

ABSTRACT

In this paper, we present a fast algorithm for joint estimation of the two-dimensional (2-D) directions of arrival (DOAs) and frequencies of the incoming signals in wireless communications using a hierarchical space–time decomposition (HSTD) technique. Based on the HSTD, the proposed algorithm makes use of a sequence of one-dimensional (1-D) Unitary estimation of signal parameters via rotational invariance technique (ESPRIT) algorithms to estimate these parameters alternatively in a hierarchical tree structure. Also, in between every other 1-D Unitary ESPRIT algorithm, a temporal filtering process or a spatial beamforming process is invoked to partition the incoming signals into finer groups stage by stage to enhance the estimation accuracy as well as to alleviate the contaminated noise. Furthermore, with such a tree-structured estimation scheme, the pairing of these parameters is automatically determined without extra computational overhead. Simulation results show that the new algorithm provides satisfactory performance but with drastically reduced computations compared with previous works.

1. Introduction

Joint estimation of the two-dimensional (2-D) directions of arrival (DOA) (azimuth and elevation angles) and carrier frequencies of multiple sources impinging on an antenna array is of importance in various facets of signal processing disciplines such as radar, sonar, radio astronomy, seismic data processing, and mobile communication systems. For example, these parameters can be applied to locating the mobiles and to allocating pilot tones in space division multiple access (SDMA) systems [1]. Also, a precise estimation of these parameters is helpful in obtaining a better channel estimate and thus enhances the system performance [2,3]. Therefore, there has been a flurry of interest for the joint 2-D DOA and frequency estimation.

The maximum likelihood (ML) approach is asymptotically efficient, but requires multidimensional search, which is in general computationally prohibitive in practice. To alleviate the computational overhead, miscellaneous high-dimensional subspace based algorithms such as MUltiple SIgnal Classification (MUSIC) or estimation of signal parameters via rotational invariance technique (ESPRIT)-based [4] algorithms have been reported to jointly estimate the 2-D DOAs and frequencies of the impinging signals. The latter are in particular computationally attractive, as they are free of a higher-dimensional search on the azimuth-elevation-frequency plane. For example, Haardt and Nossek [1] addressed a three-dimensional (3-D) Unitary ESPRIT to estimate these three parameters by extending the computationally efficient 1-D Unitary ESPRIT [5] to 3-D scenarios. Strobach [6] incorporated a total least squares phased averaging method in the 3-D ESPRIT for better signal subspace

☆ This work was supported by National Science Council of ROC under Contract NSC 97-2221-E-011-053.

* Corresponding author. Tel.: +886 2 27376412; fax: +886 2 27376424.
E-mail addresses: whf@mail.ntust.edu.tw (W.-H. Fang), jdlin@ea.njtc.edu.tw (J.-D. Lin).

0165-1684/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.sigpro.2009.06.011

estimates in order to increase the estimation accuracy of these three parameters, which is especially suitable for large arrays. Despite the effectiveness of [1,6], both of them still call for enormous computational overhead due to higher-dimensional data stacking and the related eigendecomposition. In light of this, an efficient ESPRIT-based algorithm was recently considered in [7], which constructed three space–time vectors using the data received by three parallel linear antenna arrays and then estimated these parameters based on the eigenvalues of their related covariance matrices. Yet another low-complexity method was proposed by Tayem et al. [8], which employed the efficient propagator method to estimate these parameters. Although the computational load is substantially alleviated, the algorithms addressed in [7,8], however, requires a special antenna array geometry.

In order to yield high estimation accuracy with low computational complexity, this paper proposes a fast, yet accurate algorithm for joint estimation of the 2-D DOAs and frequencies of the incoming signals based on the hierarchical space–time decomposition (HSTD) technique, which has been successfully applied to 1-D MUSIC-based DOA-delay and DOA-frequency estimation [9,10]. The essence of the HSTD scheme lies in a succinct combination of the parameter estimation and temporal filtering/spatial beamforming processes, in which the parameters are estimated alternatively in a hierarchical tree structure with signals being partitioned into finer groups stage by stage [11]. More specifically, the data in every stage are ingeniously re-arranged so that the proposed algorithm can make use of a sequence of the 1-D Unitary ESPRIT algorithms to estimate the 2-D DOAs and frequencies in a coarse-fine manner. In between every other 1-D Unitary ESPRIT, either a temporal filtering process or a spatial beamforming process is invoked to partition the incoming signals into finer groups in the parameter estimation process. The partitioning of signals, implemented by a set of projection matrices based on the resolvable parameters estimated in the previous stage, aims to enhance the estimation accuracy and to alleviate the contaminated noise. Based on such an HSTD scheme, not only accurate 2-D DOA/frequency estimates can be obtained, but the pairing of these parameters is automatically achieved without extra computations. Conducted simulation results show that the new algorithm provides satisfactory performance but with drastically reduced computations compared with previous works.

The rest of this paper is organized as follows. Section 2 introduces the signal model considered in this paper. In Section 3, we describe the proposed 1-D Unitary ESPRIT-based algorithm which hierarchically estimates the 2-D DOAs and carrier frequencies of the incoming signals. Simulation results are presented in Section 4 to verify the proposed approach. Section 5 provides a concluding remark to summarize the paper.

1.1. Mathematical notation

Throughout this paper, vectors and matrices are denoted by lower-case and upper-case boldface letters, respectively. In addition, the mathematical notations are denoted as follows: $(\cdot)^*$ = complex conjugation, $(\cdot)^T$ = matrix transposition, $(\cdot)^H$ = Hermitian operation, $E[\cdot]$ = mathematical expectation, diag$\{\mathbf{v}\}$ = diagonal matrix whose diagonal is the vector $\mathbf{v}$, $\delta(\cdot)$ = Kronecker delta function, Re$\{\mathbf{A}\}$ = real part of matrix $\mathbf{A}$, Im$\{\mathbf{A}\}$ = imaginary part of matrix $\mathbf{A}$, trace($\mathbf{A}$) = summation of diagonal elements of matrix $\mathbf{A}$, vec$(\cdot)$ = column stacking operation, $\otimes$ = Kronecker product [12], $\mathbf{A}(i:j)$ = submatrix consisting of the ith to jth columns of matrix $\mathbf{A}$, $\mathbf{A}(i)$ = ith column of matrix $\mathbf{A}$, $\mathbf{A}(i,j)$ = (i,j)th element of matrix $\mathbf{A}$, $\mathbf{a}(i)$ = ith element of vector $\mathbf{a}$, and $\|\cdot\|$ = vector 2-norm.

2. Signal model

Consider a uniform rectangular array (URA) with $M \times N$ omni-directional antennas, which are placed on the $x-y$ plane and the interval distances of antennas in the x-axis and y-axis are dx and dy, respectively. Each antenna is followed by a tapped delay line with L time delay elements of delay T_s, as shown in Fig. 1. Assume that there are K uncorrelated narrowband sources $\{s_k(t)\}$, each of which is carried by the frequency f_k, impinging on the URA.

By sampling the output of each antenna at a rate $f_s = 1/T_s$, the observed signal at time t at the lth delay element output of the (m,n)th antenna element can be expressed as

$$x_{mnl}(t) = \sum_{k=1}^{K} s_k(t)\, e^{-j2\pi(((M+1)/2)-m)u_k}\, e^{-j2\pi(((N+1)/2)-n)v_k} \times e^{-j2\pi(((L+1)/2)-l)f_k T_s} + n_{mnl}(t) \quad (1)$$

for $m = 1,\ldots,M, n = 1,\ldots,N, l = 1,\ldots,L$, where $u_k = (dx \cdot f_k/c)\sin\phi_k \cos\theta_k$ and $v_k = (dy \cdot f_k/c)\sin\phi_k \sin\theta_k$, in which ϕ_k and θ_k are the elevation and azimuth angles of the kth signal, respectively, and c is the wave propagation speed. $n_{mnl}(t)$ denotes the additive zero-mean complex white Gaussian noise with variance σ^2 at the lth delay element of the (m,n)th antenna. For simplicity, assume that the antenna spacing dx and dy are both equal to half of the minimum wavelength of the impinging signal sources. We also assume that the impinging signals have at least one distinct component in (u,v,f). Also note that for the

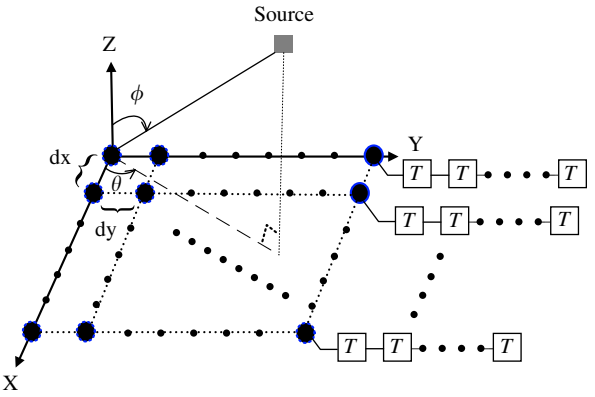

Fig. 1. Uniform rectangular array with tapped delay lines.

employment of the Unitary ESPRIT to be discussed next, the reference point is set at the center of the URA and, without loss of generality, M, N, and L are all assumed to be even.

3. Proposed 1-D Unitary ESPRIT-based algorithm

The proposed algorithm begins with the estimation of the carrier frequencies of the incoming signals, as in general we have more temporal data to render a precise frequency estimate. To achieve this, we first construct $\mathbf{X}_f$ by stacking $x_{mnl}(t)$ as

$$\mathbf{X}_f = [\mathbf{x}_{11}^{\swarrow}(t), \ldots, \mathbf{x}_{M1}^{\swarrow}(t) \vdots \cdots \vdots \mathbf{x}_{1N}^{\swarrow}(t), \ldots, \mathbf{x}_{MN}^{\swarrow}(t)] \tag{2}$$

where the time index t in $\mathbf{X}_f$ is omitted for notational brevity and $\mathbf{x}_{mn}^{\swarrow}(t) = [x_{mn1}(t), \ldots, x_{mnL}(t)]^T$, $m = 1, \ldots, M$, $n = 1, \ldots, N$. Based on (1), it can be readily shown that

$$\begin{aligned}\mathbf{x}_{mn}^{\swarrow}(t) = &\sum_{k=1}^{K} s_k(t)\, e^{-j2\pi(((M+1)/2)-m)u_k}\, e^{-j2\pi(((N+1)/2)-n)v_k} \\ &\times [e^{-j2\pi((L-1)/2)f_k T_s}, \ldots, e^{j2\pi((L-1)/2)f_k T_s}]^T \\ &+ [n_{mn1}(t), \ldots, n_{mnL}(t)]^T, \\ &m = 1, \ldots, M, \;\; n = 1, \ldots, N\end{aligned} \tag{3}$$

As such, it follows that (2) can be re-expressed as

$$\begin{aligned}\mathbf{X}_f &= \left[\sum_{k=1}^{K} s_k(t)\mathbf{v}_k(1)\mathbf{g}_k\mathbf{u}_k^T \vdots \cdots \vdots \sum_{k=1}^{K} s_k(t)\mathbf{v}_k(N)\mathbf{g}_k\mathbf{u}_k^T\right] + \mathbf{N}_f \\ &= \sum_{k=1}^{K} s_k(t)\mathbf{g}_k(\mathbf{v}_k \otimes \mathbf{u}_k)^T + \mathbf{N}_f\end{aligned} \tag{4}$$

where $\mathbf{g}_k = [e^{-j2\pi((L-1)/2)f_k T_s}, \ldots, e^{j2\pi((L-1)/2)f_k T_s}]^T$, $\mathbf{u}_k = [e^{-j2\pi((M-1)/2)u_k}, \ldots, e^{j2\pi((M-1)/2)u_k}]^T$, and $\mathbf{v}_k = [e^{-j2\pi((N-1)/2)v_k}, \ldots, e^{j2\pi((N-1)/2)v_k}]^T$. Also, the noise matrix $\mathbf{N}_f$ is constructed by $\{n_{mnl}(t)\}$ in the same way as $\mathbf{X}_f$ from $\{x_{mnl}(t)\}$ given in (2).

We then consider the *frequency covariance matrix* of $\mathbf{X}_f$, $\mathbf{R}_f \triangleq (1/MN)E[\mathbf{X}_f\mathbf{X}_f^H]$, which, based on (4), can be shown as

$$\begin{aligned}\mathbf{R}_f &= \frac{1}{MN}E\left[\left(\sum_{k=1}^{K} s_k(t)\mathbf{g}_k(\mathbf{v}_k \otimes \mathbf{u}_k)^T + \mathbf{N}_f\right)\right. \\ &\quad \left.\times\left(\sum_{l=1}^{K} s_l(t)\mathbf{g}_l(\mathbf{v}_l \otimes \mathbf{u}_l)^T + \mathbf{N}_f\right)^H\right] \\ &= \mathbf{G}\boldsymbol{\Lambda}\mathbf{G}^H + \sigma^2\mathbf{I}\end{aligned} \tag{5}$$

where we have used the facts that $(\mathbf{v}_k \otimes \mathbf{u}_k)^T(\mathbf{v}_k \otimes \mathbf{u}_k)^* = MN$ [13] and that the noise $n_{mnl}(t)$ is white and independent of the impinging signals. $\mathbf{G} = [\mathbf{g}_1, \ldots, \mathbf{g}_K]$ is the frequency signature matrix and $\boldsymbol{\Lambda} = E[\mathbf{S}(t)\mathbf{S}^H(t)]$, in which $\mathbf{S}(t) = \mathrm{diag}\{s_1(t), \ldots, s_K(t)\}$. Note that $\mathbf{R}_f$ and $\mathbf{G}$ share the same column space and thus the 1-D Unitary ESPRIT can be employed to estimate the frequencies. More specifically, consider a transformed covariance matrix of $\mathbf{R}_f$ given by

$$\mathbf{S}_{R_f} = \mathrm{Re}\{\mathbf{Q}_L^H\mathbf{R}_f\mathbf{Q}_L\} \tag{6}$$

where

$$\mathbf{Q}_L = \frac{1}{\sqrt{2}}\begin{bmatrix}\mathbf{I}_{L/2} & j\mathbf{I}_{L/2} \\ \mathbf{J}_{L/2} & -j\mathbf{J}_{L/2}\end{bmatrix} \tag{7}$$

in which $\mathbf{I}_k$ and $\mathbf{J}_k$ are $k \times k$ identity matrix and exchange matrix, respectively.

Taking the eigendecomposition of $\mathbf{S}_{R_f}$ and using the rotational invariance property [4] results in

$$\mathbf{K}_{f1}\mathbf{E}_f\boldsymbol{\Psi}_f = \mathbf{K}_{f2}\mathbf{E}_f \tag{8}$$

where $\mathbf{E}_f$ is constructed by the eigenvectors corresponding to the K largest eigenvalues of $\mathbf{S}_{R_f}$, $\boldsymbol{\Psi}_f = \mathrm{diag}\{\tan(\pi f_1 T_s), \ldots, \tan(\pi f_K T_s)\}$, $\mathbf{K}_{f1} = \mathrm{Re}\{\mathbf{Q}_{L-1}^H\mathbf{J}_{f2}\mathbf{Q}_L\}$, and $\mathbf{K}_{f2} = \mathrm{Im}\{\mathbf{Q}_{L-1}^H\mathbf{J}_{f2}\mathbf{Q}_L\}$, in which $\mathbf{J}_{f2} = [\mathbf{0}_{(L-1)\times 1} \vdots \mathbf{I}_{L-1}]$. Using the total least squares method, we then have $\hat{\boldsymbol{\Psi}}_f = -\mathbf{V}_{12}\mathbf{V}_{22}^{-1}$, where $\mathbf{V}_{12}$ and $\mathbf{V}_{22}$ are the (1,2) and (2,2) $K \times K$ sub-matrices of $\hat{\mathbf{C}}$, respectively, in which $\hat{\mathbf{C}} = [\mathbf{K}_{f1}\mathbf{E}_f \vdots \mathbf{K}_{f2}\mathbf{E}_f]^T[\mathbf{K}_{f1}\mathbf{E}_f \vdots \mathbf{K}_{f2}\mathbf{E}_f]$. The frequencies can then be determined by taking the arctangent of the K largest eigenvalues of $\mathrm{Re}\{\hat{\boldsymbol{\Psi}}_f\}$ [5].

However, the ESPRIT algorithms cannot well resolve the parameters when they are close to each other [14], as the data matrix will be rank deficient and the related signature matrix tends to be ill-conditioned. To overcome this setback, we use the HSTD technique addressed by partitioning the signals into smaller groups based on the resolvable carrier frequencies estimated above before proceeding to estimate the 2-D DOAs. To achieve this, suppose that after carrying out the 1-D Unitary ESPRIT with respect to $\mathbf{R}_f$, we obtain a set of coarse frequency estimates, say, $\{\hat{f}_1, \ldots, \hat{f}_q\}$, where q is the number of resolvable frequencies. Based on these estimated parameters, we can construct a set of *temporal* projection matrices $\mathbf{P}_{f_i}$ given by

$$\mathbf{P}_{f_i} = \mathbf{I} - \bar{\mathbf{G}}_i(\bar{\mathbf{G}}_i^H\bar{\mathbf{G}}_i)^{-1}\bar{\mathbf{G}}_i^H, \quad i = 1, \ldots, q \tag{9}$$

where $\bar{\mathbf{G}}_i = [\hat{\mathbf{g}}_1, \ldots, \hat{\mathbf{g}}_{i-1}, \hat{\mathbf{g}}_{i+1}, \ldots, \hat{\mathbf{g}}_q]$, in which $\hat{\mathbf{g}}_k$ is obtained by replacing the true frequency f by the estimated one, $\hat{f}$, in $\mathbf{g}_k$. We then use these projection matrices to obtain a set of filtered data matrices $\mathbf{X}_{f_i} = \mathbf{P}_{f_i}\mathbf{X}_f$, $i = 1, \ldots, q$, which, based on the data model in (4), can be re-written as

$$\begin{aligned}\mathbf{X}_{f_i} &= \sum_{k=1}^{K} s_k(t)\mathbf{P}_{f_i}\mathbf{g}_k(\mathbf{v}_k \otimes \mathbf{u}_k)^T + \mathbf{P}_{f_i}\mathbf{N}_f \\ &\cong \sum_{j=1}^{r_i} s_{i,j}(t)\bar{\mathbf{g}}_{i,j}(\mathbf{v}_{i,j} \otimes \mathbf{u}_{i,j})^T + \mathbf{N}_{f_i}, \quad i = 1, \ldots, q\end{aligned} \tag{10}$$

where $\bar{\mathbf{g}}_{i,j} = \mathbf{P}_{f_i}\mathbf{g}_{i,j}$ and $\mathbf{N}_{f_i} = \mathbf{P}_{f_i}\mathbf{N}_f$, and we have used the fact that $\|\mathbf{P}_{f_i}\mathbf{g}_{l,j}\| \approx 0$, $l = 1, \ldots, q$, $l \neq i$, i.e. the incoming sources except those in the ith group will be approximately annihilated by such a temporal filtering process. Consequently, the signals are now partitioned into q groups with each group having r_i signals with close frequencies (close to $\hat{f}_i$). It is noteworthy that the filtered data matrix $\mathbf{X}_{f_i}$ only contains signals with close frequencies but well-separated (u, v)'s. Thereby, the signals with close 2-D DOAs are partitioned into different groups by assumption and can then be easily resolved.

Next, in order to estimate u by using the 1-D Unitary ESPRIT based on the filtered data $\mathbf{X}_{f_i}$, we partition $\mathbf{X}_{f_i}$ columnwise into $NL\times M$ sub-block matrices and then stack them into $\mathbf{X}_{u_i}$ as

$$\mathbf{X}_{u_i}=[(\mathbf{X}_{f_i}(1:M))^T\vdots\cdots\vdots(\mathbf{X}_{f_i}((N-1)M+1:NM))^T],\quad i=1,\ldots,q \tag{11}$$

By invoking (10), Eq. (11) can be re-expressed as

$$\begin{aligned}\mathbf{X}_{u_i}&\cong\left[\left(\sum_{j=1}^{r_i}s_{i,j}(t)\mathbf{v}_{i,j}(1)\bar{\mathbf{g}}_{i,j}\mathbf{u}_{i,j}^T+\mathbf{N}_{f_i}(1:M)\right)^T\right.\\&\quad\left.\vdots\cdots\vdots\left(\sum_{j=1}^{r_i}s_{i,j}(t)\mathbf{v}_{i,j}(N)\bar{\mathbf{g}}_{i,j}\mathbf{u}_{i,j}^T+\mathbf{N}_{f_i}((N-1)M+1:NM)\right)^T\right]\\&=\sum_{j=1}^{r_i}s_{i,j}(t)\mathbf{u}_{i,j}[\mathbf{v}_{i,j}(1)\bar{\mathbf{g}}_{i,j}^T\vdots\cdots\vdots\mathbf{v}_{i,j}(N)\bar{\mathbf{g}}_{i,j}^T]+\mathbf{N}_{u_i}\\&=\sum_{j=1}^{r_i}s_{i,j}(t)\mathbf{u}_{i,j}(\mathbf{v}_{i,j}\otimes\bar{\mathbf{g}}_{i,j})^T+\mathbf{N}_{u_i},\quad i=1,\ldots,q\end{aligned} \tag{12}$$

where we have used the fact $(\mathbf{v}_{i,j}\otimes\bar{\mathbf{g}}_{i,j})^T=[\mathbf{v}_{i,j}(1)\bar{\mathbf{g}}_{i,j}^T\vdots\cdots\vdots\mathbf{v}_{i,j}(N)\bar{\mathbf{g}}_{i,j}^T]$. Also, $\mathbf{N}_{u_i}$ is constructed by $\mathbf{N}_{f_i}$ in the same way as $\mathbf{X}_{u_i}$ from $\mathbf{X}_{f_i}$ in (11). It is shown in Appendix A that the covariance matrix of $\mathbf{N}_{u_i}$ is given by

$$\frac{1}{LN}E[\mathbf{N}_{u_i}\mathbf{N}_{u_i}^H]=\sigma_i^2\mathbf{I} \tag{13}$$

where $\sigma_i^2=((L-q+1)/L)\sigma^2$ is the new noise power. Eq. (13) implies that the noise components in $\mathbf{N}_{u_i}$ remain white and that the noise power is reduced after the temporal filtering process.

We can then compute the *spatial* covariance matrices of $\mathbf{X}_{u_i}$, $\mathbf{R}_{u_i}\triangleq(1/LN)E[\mathbf{X}_{u_i}\mathbf{X}_{u_i}^H]$, which, based on (12), can be readily shown to be

$$\mathbf{R}_{u_i}=\mathbf{U}_i\mathbf{\Lambda}_i\mathbf{U}_i^H+\sigma_i^2\mathbf{I},\quad i=1,\ldots,q \tag{14}$$

where we have used the fact that $(\mathbf{v}_{i,j}\otimes\bar{\mathbf{g}}_{i,j})^T(\mathbf{v}_{i,j}\otimes\bar{\mathbf{g}}_{i,j})^*=LN$. $\mathbf{U}_i=[\mathbf{u}_{i,1},\ldots,\mathbf{u}_{i,r_i}]$ is the signature matrix of the signals in the ith group and $\mathbf{\Lambda}_i=E[\mathbf{S}_i(t)\mathbf{S}_i^H(t)]$, in which $\mathbf{S}_i(t)=\mathrm{diag}\{s_{i,1}(t),\ldots,s_{i,r_i}(t)\}$. Carrying out the 1-D unitary ESPRIT again for each group, we can get a set of estimates of u, say, $\{\hat{u}_{i,j}\}$, $i=1,2,\ldots,q$, $j=1,2,\ldots,\rho_i$, where ρ_i is the number of resolvable u's in the ith group. Thereafter, we use these estimates to construct a set of *spatial* projection matrices given by

$$\mathbf{P}_{u_{i,j}}=\mathbf{I}-\bar{\mathbf{U}}_{i,j}(\bar{\mathbf{U}}_{i,j}^H\bar{\mathbf{U}}_{i,j})^{-1}\bar{\mathbf{U}}_{i,j}^H,\quad i=1,\ldots,q,\ j=1,\ldots,\rho_i \tag{15}$$

where $\bar{\mathbf{U}}_{i,j}=[\hat{\mathbf{u}}_{i,1}\ \ldots\ \hat{\mathbf{u}}_{i,j-1}\ \hat{\mathbf{u}}_{i,j+1}\ \ldots\ \hat{\mathbf{u}}_{i,\rho_i}]$, in which $\hat{\mathbf{u}}_{i,j}$ is to replace the true u by the estimated one in $\mathbf{u}_{i,j}$. Pre-multiplying $\mathbf{X}_{u_i}$ by $\mathbf{P}_{u_{i,j}}$ yields a set of finer group of signals $\mathbf{X}_{u_{i,j}}=\mathbf{P}_{u_{i,j}}\mathbf{X}_{u_i}$, which, using (12), can be expressed as

$$\begin{aligned}\mathbf{X}_{u_{i,j}}&=\sum_{j=1}^{\rho_i}s_{i,j}(t)\mathbf{P}_{u_{i,j}}\mathbf{u}_{i,j}(\mathbf{v}_{i,j}\otimes\bar{\mathbf{g}}_{i,j})^T+\mathbf{P}_{u_{i,j}}\mathbf{N}_{u_i}\\&\cong\sum_{l=1}^{z_{i,j}}s_{i,j,l}(t)\bar{\mathbf{u}}_{i,j,l}(\mathbf{v}_{i,j,l}\otimes\bar{\mathbf{g}}_{i,j,l})^T+\mathbf{N}_{u_{i,j}},\\&\quad i=1,\ldots,q,\ j=1,\ldots,\rho_i\end{aligned} \tag{16}$$

where we have partitioned the signals in the ith group into ρ_i subgroups with each subgroup having $z_{i,j}$ signals with close u components (close to $\hat{u}_{i,j}$), $i=1,\ldots,q$, $j=1,\ldots,\rho_i$, and used the fact that $\|\mathbf{P}_{u_{i,j}}\mathbf{u}_{i,\tau,l}\|\approx 0$, $\tau=1,\ldots,\rho_i$, $\tau\neq j$, i.e. the incoming sources except those in the (i,j)th subgroup are approximately eliminated by $\mathbf{P}_{u_{i,j}}$, and $\mathbf{N}_{u_{i,j}}=\mathbf{P}_{u_{i,j}}\mathbf{N}_{u_i}$.

We can note that the incoming signals in the (i,j)th subgroup, which possess close u components, will have diverse v's as assumed. As such, the v components for the signals in each subgroup can be well resolved. To estimate v by the 1-D Unitary ESPRIT, we partition the filtered data matrix $\mathbf{X}_{u_{i,j}}$ in (16) columnwise into $NM\times L$ sub-block matrices and then stack them as

$$\mathbf{X}_{v_{i,j}}=[\mathrm{vec}\{(\mathbf{X}_{u_{i,j}}(1:L))^T\}\vdots\cdots\vdots\mathrm{vec}\{(\mathbf{X}_{u_{i,j}}((N-1)L+1:NL))^T\}]^T \tag{17}$$

which, based on (16), can be re-expressed as

$$\begin{aligned}\mathbf{X}_{v_{i,j}}&\cong\left[\mathrm{vec}\left\{\sum_{l=1}^{z_{i,j}}s_{i,j,l}(t)\mathbf{v}_{i,j,l}(1)\bar{\mathbf{g}}_{i,j,l}\bar{\mathbf{u}}_{i,j,l}^T+(\mathbf{N}_{u_{i,j}}(1:L))^T\right\}\right.\\&\quad\vdots\cdots\vdots\left.\mathrm{vec}\left\{\sum_{l=1}^{z_{i,j}}s_{i,j,l}(t)\mathbf{v}_{i,j,l}(N)\bar{\mathbf{g}}_{i,j,l}\bar{\mathbf{u}}_{i,j,l}^T+(\mathbf{N}_{u_{i,j}}((N-1)L+1:NL))^T\right\}\right]^T\\&=\sum_{l=1}^{z_{i,j}}s_{i,j,l}(t)[\mathbf{v}_{i,j,l}(1)\bar{\mathbf{u}}_{i,j,l}\otimes\bar{\mathbf{g}}_{i,j,l}\vdots\cdots\vdots\mathbf{v}_{i,j,l}(N)\bar{\mathbf{u}}_{i,j,l}\otimes\bar{\mathbf{g}}_{i,j,l}]^T+\mathbf{N}_{v_{i,j}}\\&=\sum_{l=1}^{z_{i,j}}s_{i,j,l}(t)\mathbf{v}_{i,j,l}(\bar{\mathbf{u}}_{i,j,l}\otimes\bar{\mathbf{g}}_{i,j,l})^T+\mathbf{N}_{v_{i,j}},\\&\quad i=1,\ldots,q,\ j=1,\ldots,\rho_i\end{aligned} \tag{18}$$

where $\mathbf{N}_{v_{i,j}}$ is obtained from $\mathbf{N}_{u_{i,j}}$ in the same way as $\mathbf{X}_{v_{i,j}}$ from $\mathbf{X}_{u_{i,j}}$ in (17) and we have used the fact $\mathrm{vec}\{\bar{\mathbf{g}}_{i,j,l}\bar{\mathbf{u}}_{i,j,l}^T\}=\bar{\mathbf{u}}_{i,j,l}\otimes\bar{\mathbf{g}}_{i,j,l}$. The covariance matrix of $\mathbf{N}_{v_{i,j}}$, $\mathbf{R}_{n_{v_{i,j}}}\triangleq(1/LM)E\{\mathbf{N}_{v_{i,j}}\mathbf{N}_{v_{i,j}}^H\}$, as shown in Appendix A, is given by

$$\frac{1}{LM}E[\mathbf{N}_{v_{i,j}}\mathbf{N}_{v_{i,j}}^H]=\sigma_{i,j}^2\mathbf{I},\quad i=1,\ldots,q,\ j=1,\ldots,\rho_i \tag{19}$$

where $\sigma_{i,j}^2=((L-q+1)(M-\rho_i+1)/LM)\sigma^2$. Eq. (19) implies that the noise components in $\mathbf{N}_{v_{i,j}}$ still remain white and that the noise power is further reduced after the spatial beamforming process. Next, we determine the covariance matrix of $\mathbf{X}_{v_{i,j}}$, $\mathbf{R}_{v_{i,j}}\triangleq(1/LM)E[\mathbf{X}_{v_{i,j}}\mathbf{X}_{v_{i,j}}^H]$, which, based on (18), can be shown to be

$$\mathbf{R}_{v_{i,j}}=\mathbf{V}_{i,j}\mathbf{\Lambda}_{i,j}\mathbf{V}_{i,j}^H+\sigma_{i,j}^2\mathbf{I},\quad i=1,\ldots,q,\ j=1,\ldots,\rho_i \tag{20}$$

where we have used the fact that $(\bar{\mathbf{u}}_{i,j} \otimes \bar{\mathbf{g}}_{i,j})^T(\bar{\mathbf{u}}_{i,j} \otimes \bar{\mathbf{g}}_{i,j})^* = LM$. $\mathbf{V}_{i,j} = [\mathbf{v}_{i,j,1}, \ldots, \mathbf{v}_{i,j,z_{i,j}}]$ is the signature matrix of $\mathbf{v}_{i,j,l}$ and $\mathbf{\Lambda}_{i,j} = E[\mathbf{S}_{i,j}(t)\mathbf{S}_{i,j}^H(t)]$, in which $\mathbf{S}_{i,j}(t) = \mathrm{diag}\{s_{i,j,1}(t), \ldots, s_{i,j,z_{i,j}}(t)\}$. Along the same line as above, the 1-D Unitary ESPRIT can be applied to $\mathbf{R}_{v_{i,j}}$ to estimate v and get, say, $\{\hat{\mathbf{v}}_{i,j,l}\}$, $i = 1, 2, \ldots, q$, $j = 1, 2, \ldots, \rho_i$, $l = 1, 2, \ldots, z_{i,j}$.

We can note that the f's and u's estimated in the previous stages are rather rough when these parameters are closely spaced. Consequently, we can get a more precise estimate of f's and u's by carrying out the 1-D Unitary ESPRIT again based on the finer groups of data $\mathbf{X}_{v_{i,j}}$. For this, we construct another set of spatial projection matrices given by

$$\mathbf{P}_{v_{i,j,l}} = \mathbf{I} - \bar{\mathbf{V}}_{i,j,l}(\bar{\mathbf{V}}_{i,j,l}^H\bar{\mathbf{V}}_{i,j,l})^{-1}\bar{\mathbf{V}}_{i,j,l}^H, \quad i = 1, \ldots, q,\ j = 1, \ldots, \rho_i,\ l = 1, \ldots, z_{i,j} \tag{21}$$

where $\bar{\mathbf{V}}_{i,j,l} = [\hat{\mathbf{v}}_{i,j,1} \ \ldots \ \hat{\mathbf{v}}_{i,j,l-1}\ \hat{\mathbf{v}}_{i,j,l+1} \ \ldots \ \hat{\mathbf{v}}_{i,j,z_{i,j}}]$, and then pre-multiply $\mathbf{X}_{v_{i,j}}$ by the projection matrix $\mathbf{P}_{v_{i,j,l}}$ to render $\mathbf{X}_{v_{i,j,l}} = \mathbf{P}_{v_{i,j,l}}\mathbf{X}_{v_{i,j}}$, which will annihilate the impinging signals that do not belong to the (i,j,l)th subgroup. Based on (18), the new data matrix can be expressed as

$$\begin{aligned}\mathbf{X}_{v_{i,j,l}} &= \sum_{m=1}^{z_{i,j}} s_{i,j,m}(t)\mathbf{P}_{v_{i,j,l}}\mathbf{v}_{i,j,m}(\bar{\mathbf{u}}_{i,j,m} \otimes \bar{\mathbf{g}}_{i,j,m})^T + \mathbf{P}_{v_{i,j,l}}\mathbf{N}_{v_{i,j}} \\ &\cong s_{i,j,l}(t)\bar{\mathbf{v}}_{i,j,l}(\bar{\mathbf{u}}_{i,j,l} \otimes \bar{\mathbf{g}}_{i,j,l})^T + \mathbf{N}_{v_{i,j,l}}\end{aligned} \tag{22}$$

for $i = 1, \ldots, q$, $j = 1, \ldots, \rho_i$, $l = 1, \ldots, z_{i,j}$, where $\bar{\mathbf{v}}_{i,j,l} = \mathbf{P}_{v_{i,j,l}}\mathbf{v}_{i,j,l}$, $\bar{\mathbf{u}}_{i,j,l} = \mathbf{P}_{u_{i,j}}\mathbf{u}_{i,j,l}$, $\bar{\mathbf{g}}_{i,j,l} = \mathbf{P}_{f_i}\mathbf{g}_{i,j,l}$, $\mathbf{N}_{v_{i,j,l}} = \mathbf{P}_{v_{i,j,l}}\mathbf{N}_{v_{i,j}}$, and we have used the fact that $\|\mathbf{P}_{v_{i,j,l}}\mathbf{v}_{i,j,\tau}\| \approx 0$, $\tau = 1, \ldots, z_{i,j}$, $\tau \neq l$. It is noteworthy that there is only one signal left in $\mathbf{X}_{v_{i,j,l}}$ after the temporal filtering process and spatial beamforming processes. This is due to the fact that, as asserted above, each group of the filtered data matrix $\mathbf{X}_{f_i}$, $i = 1, \ldots, q$, contains signals with well-separated (u,v)'s, which implies that at least one component in (u,v) is diverse. Now since each (i,j)th subgroup of data, $i = 1, \ldots, q$, $j = 1, \ldots, \rho_i$, has close u components, the signals inside this subgroup must have diverse v components and can then be well resolved. Therefore, after the spatial beamforming process addressed above, the signals in the (i,j)th subgroup are partitioned into $z_{i,j}$ finer subgroups of data with each subgroup containing only one signal. We can then get finer estimations of u and f based on these finer groups of data.

More specifically, for a more accurate estimation of f, we partition the filtered data matrix $\mathbf{X}_{v_{i,j,l}}$ into $MN \times L$ sub-block matrices and then stack them as

$$\begin{aligned}\mathbf{X}_{f_{i,j,l}} &= [(\mathbf{X}_{v_{i,j,l}}(1:L))^T \vdots \cdots \vdots (\mathbf{X}_{v_{i,j,l}}((M-1)L+1:ML))^T] \\ &\cong [s_{i,j,l}(t)\bar{\mathbf{u}}_{i,j,l}(1)\bar{\mathbf{g}}_{i,j,l}\bar{\mathbf{v}}_{i,j,l}^T + (\mathbf{N}_{v_{i,j,l}}(1:L))^T \\ &\quad \vdots \cdots \vdots s_{i,j,l}(t)\bar{\mathbf{u}}_{i,j,l}(M)\bar{\mathbf{g}}_{i,j,l}\bar{\mathbf{v}}_{i,j,l}^T \\ &\quad + (\mathbf{N}_{v_{i,j,l}}((M-1)L+1:ML))^T] \\ &= s_{i,j,l}(t)\bar{\mathbf{g}}_{i,j,l}(\bar{\mathbf{u}}_{i,j,l} \otimes \bar{\mathbf{v}}_{i,j,l})^T + \mathbf{N}_{f_{i,j,l}}\end{aligned} \tag{23}$$

for $i = 1, \ldots, q$, $j = 1, \ldots, \rho_i$, $l = 1, \ldots, z_{i,j}$, where we have used the fact that $(\bar{\mathbf{u}}_{i,j,l} \otimes \bar{\mathbf{v}}_{i,j,l})^T = [\bar{\mathbf{v}}_{i,j,l}^T\bar{\mathbf{u}}_{i,j,l}(1) \vdots \cdots \vdots \bar{\mathbf{v}}_{i,j,l}^T\bar{\mathbf{u}}_{i,j,l}(M)]$ and $\mathbf{N}_{f_{i,j,l}}$ is obtained from $\mathbf{N}_{v_{i,j,l}}$ in the same way as $\mathbf{X}_{f_{i,j,l}}$ from $\mathbf{X}_{v_{i,j,l}}$ in (23). Again, it is shown in Appendix A that the covariance matrix of $\mathbf{N}_{f_{i,j,l}}$ is given by

$$\frac{1}{MN}E[\mathbf{N}_{f_{i,j,l}}\mathbf{N}_{f_{i,j,l}}^H] = \sigma_{f_{i,j,l}}^2\mathbf{P}_{f_i}, \quad i = 1, \ldots, q,\ j = 1, \ldots, \rho_i,\ l = 1, \ldots, z_{i,j} \tag{24}$$

where $\mathbf{P}_{f_i}$ is as defined in (9) and $\sigma_{f_{i,j,l}}^2 = ((M-\rho_i+1)(N-z_{i,j}+1)/MN)\sigma^2$. By using (24), the covariance matrix of $\mathbf{X}_{f_{i,j,l}}$, $\mathbf{R}_{f_{i,j,l}} \triangleq (1/MN)\mathbf{E}[\mathbf{X}_{f_{i,j,l}}\mathbf{X}_{f_{i,j,l}}^H]$, can be re-written as

$$\mathbf{R}_{f_{i,j,l}} = \xi_{i,j,l}^2\bar{\mathbf{g}}_{i,j,l}\bar{\mathbf{g}}_{i,j,l}^H + \sigma_{f_{i,j,l}}^2\mathbf{P}_{f_i}, \quad i = 1, \ldots, q,\ j = 1, \ldots, \rho_i,\ l = 1, \ldots, z_{i,j} \tag{25}$$

where $\xi_{i,j,l}^2 = E[s_{i,j,l}(t)s_{i,j,l}^*(t)]$ and we have used fact that $(\bar{\mathbf{v}}_{i,j,l} \otimes \bar{\mathbf{u}}_{i,j,l})^T(\bar{\mathbf{v}}_{i,j,l} \otimes \bar{\mathbf{u}}_{i,j,l})^* = MN$. We can notice that $\bar{\mathbf{g}}_{i,j,l} = \mathbf{P}_{f_i}\mathbf{g}_{i,j,l}$ in (25) does not possess the Vandermonde structure, as $\mathbf{g}_{i,j,l}$ has been pre-multiplied by the projection matrix $\mathbf{P}_{f_i}$ and thus the Unitary ESPRIT is not applicable. Moreover, the noise components in $\mathbf{N}_{f_{i,j,l}}$ are now no longer white, as shown in (24). To overcome these setbacks, we utilize the fact that $\bar{\mathbf{g}}_{i,j,l}$ in (23) is the eigenvector corresponds to the largest eigenvalue of $\mathbf{R}_{f_{i,j,l}}$, as shown in Appendix B. As such, $\bar{\mathbf{g}}_{i,j,l}$ is orthogonal to $\mathbf{I} - \mathbf{e}_{f_{i,j,l}}\mathbf{e}_{f_{i,j,l}}^H$, where $\mathbf{e}_{f_{i,j,l}} = \bar{\mathbf{g}}_{f_{i,j,l}}/\|\bar{\mathbf{g}}_{f_{i,j,l}}\|$, or equivalently,

$$(\mathbf{I} - \mathbf{e}_{f_{i,j,l}}\mathbf{e}_{f_{i,j,l}}^H)\mathbf{P}_{f_i}\mathbf{g}_{i,j,l} = \mathbf{0}, \quad i = 1, \ldots, q,\ j = 1, \ldots, \rho_i,\ l = 1, \ldots, z_{i,j} \tag{26}$$

where we have used the definition that $\bar{\mathbf{g}}_{i,j,l} = \mathbf{P}_{f_i}\mathbf{g}_{i,j,l}$. Therefore, $\mathbf{g}_{i,j,l}$ belongs to the subspace spanned by the column space of $\mathbf{\Xi}_{f_{i,j,l}}$, where $\mathbf{\Xi}_{f_{i,j,l}} = \mathbf{I} - (\mathbf{I} - \mathbf{e}_{f_{i,j,l}}\mathbf{e}_{f_{i,j,l}}^H)\mathbf{P}_{f_i}$. Denote $\mathbf{G}_{i,j,l} = [\bar{\mathbf{G}}_i\ \mathbf{G}_{i,j,l}]$, where $\bar{\mathbf{G}}_i$ is as defined in (9). It is shown in Appendix C that

$$\mathbf{\Xi}_{f_{i,j,l}} = \mathbf{G}_{i,j,l}(\mathbf{G}_{i,j,l}^H\mathbf{G}_{i,j,l})^{-1}\mathbf{G}_{i,j,l}^H, \quad i = 1, \ldots, q,\ j = 1, \ldots, \rho_i,\ l = 1, \ldots, z_{i,j} \tag{27}$$

where $\mathbf{G}_{i,j,l} = [\bar{\mathbf{G}}_i\ \mathbf{g}_{i,j,l}]$. Note that the former of $\mathbf{G}_{i,j,l}$ is related to the previously estimated frequencies whereas the latter is related to the frequency of the signal in the (i,j,l)th subgroup, $j = 1, \ldots, \rho_i$, $l = 1, \ldots, z_{i,j}$. Also, since $\mathbf{G}_{i,j,l}$ retains the Vandermonde structure and shares the same column space as $\mathbf{\Xi}_{f_{i,j,l}}$, the 1-D Unitary F-ESPRIT can now be utilized for $\mathbf{\Xi}_{f_{i,j,l}}$, and a more precise estimate of f's can be obtained.

Similarly, to attain a more precise estimate of u we partition the data matrix $\mathbf{X}_{v_{i,j,l}}$ into $M\ N \times L$ sub-block matrices and then stack them as

$$\begin{aligned}\mathbf{X}_{u_{i,j,l}} &= [\mathrm{vec}\{(\mathbf{X}_{v_{i,j,l}}(1:L))^T\} \vdots \cdots \vdots \mathrm{vec}\{(\mathbf{X}_{v_{i,j,l}}((M-1)L+1:ML))^T\}]^T \\ &\cong [\mathrm{vec}\{s_{i,j,l}(t)\bar{\mathbf{u}}_{i,j,l}(1)\bar{\mathbf{g}}_{i,j,l}\bar{\mathbf{v}}_{i,j,l}^T + (\mathbf{N}_{v_{i,j,l}}(1:L))^T\} \\ &\quad \vdots \cdots \vdots \mathrm{vec}\{s_{i,j,l}(t)\bar{\mathbf{u}}_{i,j,l}(M)\bar{\mathbf{g}}_{i,j,l}\bar{\mathbf{v}}_{i,j,l}^T \\ &\quad + (\mathbf{N}_{v_{i,j,l}}((M-1)L+1:ML))^T\}]^T \\ &= s_{i,j,l}(t)\bar{\mathbf{u}}_{i,j,l}(\bar{\mathbf{v}}_{i,j,l} \otimes \bar{\mathbf{g}}_{i,j,l})^T + \mathbf{N}_{u_{i,j,l}}\end{aligned} \tag{28}$$

for $i = 1, \ldots, q$, $j = 1, \ldots, \rho_i$, $l = 1, \ldots, z_{i,j}$, where $\mathbf{N}_{u_{i,j,l}}$ is obtained from $\mathbf{N}_{v_{i,j,l}}$ in the same way as $\mathbf{X}_{u_{i,j,l}}$ from $\mathbf{X}_{v_{i,j,l}}$ in (28) and we have used the fact that $\mathrm{vec}\{\bar{\mathbf{g}}_{i,j,l}\bar{\mathbf{v}}_{i,j,l}^T\} = \bar{\mathbf{v}}_{i,j,l} \otimes \bar{\mathbf{g}}_{i,j,l}$. Thereafter, working with the covariance matrix of $\mathbf{X}_{u_{i,j,l}}$, $\mathbf{R}_{u_{i,j,l}} \triangleq (1/LN)E[\mathbf{X}_{u_{i,j,l}}\mathbf{X}_{u_{i,j,l}}^H]$ and following the same steps as the above derivations results in a more precise estimate of $u_{i,j,l}$'s.

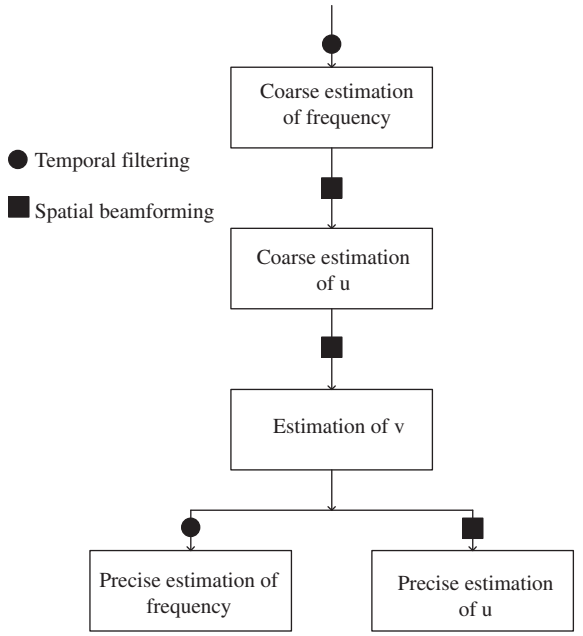

Fig. 2. The flowchart of the proposed algorithm.

For easy reference, the overall structure of the proposed hierarchical tree-structured 1-D Unitary ESPRIT based algorithm is as shown in Fig. 2 with each step being delineated in the following:

Step 1: *Rough frequency estimation*. Estimate the covariance matrix $\hat{\mathbf{R}}_f = (1/SMN)\sum_{s=1}^{S}\mathbf{X}_f(t_s)\mathbf{X}_f^H(t_s)$, where S is the number of snapshots and $\mathbf{X}_f(t_s) = [\mathbf{x}_{11}^{\swarrow}(t_s), \ldots, \mathbf{x}_{M1}^{\swarrow}(t_s)\cdots\mathbf{x}_{1N}^{\swarrow}(t_s), \ldots, \mathbf{x}_{MN}^{\swarrow}(t_s)]$, is as defined in (2), and then invoke the 1-D Unitary ESPRIT to yield a set of group frequency estimates $\{\hat{f}_1, \ldots, \hat{f}_q\}$, where q is the number of resolvable frequencies.

Step 2: *Temporal filtering*. Employ $\{\hat{f}_1, \ldots, \hat{f}_q\}$ to construct the projection matrix $\mathbf{P}_{f_i}$ by (9) and then use $\mathbf{P}_{f_i}$ to obtain the filtered data matrix $\mathbf{X}_{f_i}(t_s) = \mathbf{P}_{f_i}\mathbf{X}_f(t_s)$, $i = 1, \ldots, q$.

Step 3: *Rough estimation of u*. Partition $\mathbf{X}_{f_i}(t_s), i = 1, \ldots, q$, and stack them based on (12) to form $\mathbf{X}_{u_i}(t_s)$ and then use them to estimate the covariance matrix $\hat{\mathbf{R}}_{u_i} = (1/SLN)\sum_{s=1}^{S}\mathbf{X}_{u_i}(t_s)\mathbf{X}_{u_i}^H(t_s)$. Thereafter, use the 1-D Unitary ESPRIT to estimate the u's to get $\{\hat{u}_{i,j}\}$, $i = 1, \ldots, q$, $j = 1, \ldots, \rho_i$, where ρ_i is the number of u's resolvable in the ith group.

Step 4: *Spatial beamforming* (*I*). Employ $\{\hat{u}_{i,j}\}$ to construct the projection matrix $\mathbf{P}_{u_{i,j}}$ by (15) and then use $\mathbf{P}_{u_{i,j}}$ to obtain the filtered data matrix $\mathbf{X}_{u_{i,j}}(t_s) = \mathbf{P}_{u_{i,j}}\mathbf{X}_{u_i}(t_s)$, $i = 1, \ldots, q$, $j = 1, \ldots, \rho_i$.

Step 5: *Estimation of v*. Partition the filtered data $\mathbf{X}_{u_{i,j}}(t_s)$, stack them based on (17) to form $\mathbf{X}_{v_{i,j}}(t_s)$ and then use $\mathbf{X}_{v_{i,j}}$ to estimate the covariance matrix $\hat{\mathbf{R}}_{v_{i,j}} = (1/SLM)\sum_{s=1}^{S}\mathbf{X}_{v_{i,j}}(t_s)\mathbf{X}_{v_{i,j}}^H(t_s)$. Next, use the 1-D Unitary ESPRIT to estimate the v's to get $\{\hat{v}_{i,j,l}\}$, $i = 1, \ldots, q$, $j = 1, \ldots, \rho_i$, $l = 1, \ldots, z_{i,j}$, where $z_{i,j}$ is the number of signals in the jth subgroup of the ith group.

Step 6: *Spatial beamforming* (*II*). Employ $\{\hat{v}_{i,j,l}\}$ to construct the projection matrices $\mathbf{P}_{v_{i,j,l}}$ as given in (21), then use $\mathbf{P}_{v_{i,j,l}}$ to obtain the filtered data matrix $\mathbf{X}_{v_{i,j,l}}(t_s) = \mathbf{P}_{v_{i,j,l}}\mathbf{X}_{v_{ij}}(t_s)$, $i = 1, \ldots, q$, $j = 1, \ldots, \rho_i, l = 1, \ldots, z_{i,j}$.

Step 7: *Precise u and frequency estimation*. Partition and re-stack $\mathbf{X}_{v_{i,j,l}}(t_s)$ as (28) and (23) to obtain $\mathbf{X}_{u_{i,j,l}}(t_s)$ and $\mathbf{X}_{f_{i,j,l}}(t_s)$, respectively. Estimate the covariance matrix $\hat{\mathbf{R}}_{f_{i,j,l}} = (1/SMN)\sum_{s=1}^{S}\mathbf{X}_{f_{i,j,l}}(t_s)\mathbf{X}_{f_{i,j,l}}^H(t_s)$, and then utilize the normalized eigenvector $\hat{\mathbf{e}}_{f_{i,j,l}}$ corresponding to its largest eigenvalue and the projection matrix $\mathbf{P}_{f_i}$ to form $\mathbf{\Xi}_{f_{i,j,l}} = \mathbf{I} - (\mathbf{I} - \hat{\mathbf{e}}_{f_{i,j,l}}\hat{\mathbf{e}}_{f_{i,j,l}}^H)\mathbf{P}_{f_i}$. Thereafter, use the 1-D Unitary ESPRIT to obtain precise frequency estimates from $\mathbf{\Xi}_{f_{i,j,l}}$. Following the above procedures based on $\mathbf{X}_{u_{i,j,l}}(t_s)$ renders precise estimates of u. Note that since every subgroup in this step only contains one signal, the pairing process is automatically achieved. Finally, we can obtain the estimates of the elevation and azimuth angles by $\hat{\phi} = \sin^{-1}(c/\hat{f}d)\sqrt{\hat{u}^2 + \hat{v}^2}$ and $\hat{\theta} = \tan^{-1}\hat{v}/\hat{u}$.

Remarks. (1) In the proposed algorithm, it is assumed that the number of groups, q, the number of signals in each group, ρ_i, and the number of signals in each subgroup, $z_{i,j}$, are known or have been perfectly estimated, say by the AIC or MDL criterion addressed in [15], as required for the employment of the ESPRIT-type algorithms.

(2) The computational complexity of the proposed algorithm can be determined as follows.[1] Step 1 requires $4L^2MNS + \frac{2}{3}L^3$ real multiplications, which include $4L^2MNS$ real multiplications for the determination of $\hat{\mathbf{R}}_f$ and $\frac{2}{3}L^3$ real multiplications for a 1-D Unitary ESPRIT with respect to the $L \times L$ matrix $\hat{\mathbf{S}}_{R_f}$. In Step 2, $4L^2MNSq$ real multiplications are needed for the temporal filtering process which involves the multiplication of the$L \times MN$ matrix $\mathbf{X}_f$ and the $L \times L$ matrix $\mathbf{P}_{f_i}$ for all q groups. In Step 3, $4LM^2NSq + \frac{2}{3}M^3q$ real multiplications are required to determine $\hat{\mathbf{R}}_{u_i}$, $i = 1, \ldots, q$, and each group needs to conduct the 1-D Unitary ESPRIT with respect to an $M \times M$ matrix and thus needs $\frac{2}{3}M^3$ real multiplications. In Step 4, $4(\sum_{i=1}^{q}\rho_i)LM^2NS$ real multiplications is required for the multiplication of the $M \times M$ matrix $\mathbf{P}_{u_{i,j}}$ and the $M \times NL$ matrix $\mathbf{X}_{u_{i,j}}$, $i = 1, \ldots, q$, $j = 1, \ldots, \rho_i$. In Step 5, for each subgroup, $4LMN^2S + \frac{2}{3}N^3$ real multiplications are required to compute $\hat{\mathbf{R}}_{v_{i,j}}$ and to conduct the related 1-D Unitary ESPRIT (with respect to an $N \times N$ matrix), respectively. In Step 6, $4LMN^2SK$ real multiplications are required for the second spatial beamforming process which involves the multiplication of the $N \times N$ matrix $\mathbf{P}_{v_{i,j,l}}$ and the $N \times LM$ matrix $\mathbf{X}_{v_{i,j}}$, $i = 1, \ldots, q$, $j = 1, \ldots, \rho_i$, $l = 1, \ldots, z_{i,j}$. Finally, in Step 7, $4L^2MNSK + \frac{2}{3}L^3K + 4LM^2NSK + \frac{2}{3}M^3K$ real multi-

[1] We have used the following rules: the multiplication of an $M \times N$ matrix and an $N \times L$ matrix requires MNL multiplication, the main computational load of the 1-D Unitary ESPRIT with respect to an $N \times N$ correlation matrix is mainly dictated by its eigendecomposition and roughly calls for $\frac{2}{3}N^3$ real multiplications [13], and a complex multiplication involves four real multiplications.

plications are required to compute the $L \times L$ matrix $\mathbf{R}_{f_{i,j,l}}$ and the $M \times M$ matrices $\mathbf{R}_{u_{i,j,l}}$, and the corresponding 1-D Unitary ESPRIT with respect to these two matrices. As a whole, roughly $4(L^2MNS(K+q+1)+LM^2NS(K+q+\sum_{i=1}^{q}\rho_i)+LMN^2S(K+\sum_{i=1}^{q}\rho_i))$ real multiplications are required for the proposed algorithm. If we assume that $S \gg M, N, L > K$, the total number of real multiplications called for is then about $4LMNSK(L+M+N)$. In contrast, [1,6] need to stack the data into a higher-dimensional data matrix in order to simultaneously estimate these parameters. More specifically, [1] requires roughly $(2S+\frac{2}{3}MNL)M^2N^2L^2$ real multiplications for the determination of an $LMN \times LMN$ (real) matrix and a 3-D Unitary ESPRIT with respect to this matrix. On the other hand, Strobach [6] needs $(8S+\frac{8}{3}MNL)M^2N^2L^2$ real multiplications for the determination of an $LMN \times LMN$ (complex) correlation matrix followed by some transformations along with a 3-D Unitary ESPRIT. Therefore, [1,6] are both far more computationally expensive than the proposed algorithm. The savings in the computations will be even more pronounced for the URA of a larger size.

(3) To reduce the computing complexity, we proceed the beamspace processing in (6) by using the unitary transformation matrix $\mathbf{Q}_L$ in (7) and then taking its real part. The advantage of this process is that we can use the real arithmetic for the subsequent operations which can give a computational saving. Also note that $\mathrm{Re}\{\mathbf{Q}_L^H\mathbf{R}_f\mathbf{Q}_L\} = \mathbf{Q}_L^H\mathbf{R}_{f,fb}\mathbf{Q}_L$, where $\mathbf{R}_{f,fb} = (1/2MN)E[\mathbf{X}_f(t)\mathbf{X}_f^H(t)+\mathbf{J}\mathbf{X}_f^*(t)\mathbf{X}_f^T(t)\mathbf{J}]$ denotes the forward–backward averaged covariance matrix and we have used the fact that $\mathbf{Q}_L^H\mathbf{J} = (\mathbf{Q}^H)^*$ [12]. Therefore, we can use the forward–backward averaging in (6) and achieve exactly the same result.

(4) Note that as there is only one signal left in the estimation of ν, it is computationally advantageous to use the conventional periodogram, which can be efficiently implemented by the FFT. Note, however, that now the periodogram is not equivalent to the optimum ML estimate since, as shown in Appendix A, the noise is no longer white Gaussian. Also, the estimate of f is based on the $L \times (MN)$ matrix $\mathbf{X}_{f_{i,j,l}}$ and thus the resolution of f will depend on L. In order to render a precise estimate of f, a large L is required.

4. Simulations and discussions

Some simulations are conducted in this section to assess the proposed approach. Assume that the receiver is a 6×6 ($M = N = 6$) URA spaced half of the minimum wavelength of the impinging sources apart, where each antenna is followed by a tapped line with $L = 12$ delay elements and the sampling frequency is $f_s = 400$ MHz. Two scenarios are considered. In the first scenario, there are $K = 4$ users and the azimuth and elevation angles of the users are $[63, 29, 75, 63]^\circ$ and $[23, 44, 46, 13]^\circ$ with center frequencies [103, 103, 103, 180] MHz, respectively. This implies that $u = [0.1015, 0.3477, 0.1065, 0.1021]$ and $\nu = [0.1992, 0.1927, 0.3976, 0.2004]$. In the second scenario, $K = 5$ and the azimuth and elevation angles are $[50, 55, 25, 16, 35]^\circ$ and $[70, 45, 41, 83, 31]^\circ$ with center frequencies [110.1, 163, 110.3, 110.3, 163] MHz, respectively, which imply that $u = [0.4080, 0.4056, 0.4024, 0.6456, 0.4219]$ and $\nu = [0.4862, 0.5792, 0.1876, 0.1851, 0.2954]$. The average fading amplitudes of all signals are equal and normalized to 0 dB. $S = 200$ symbols are employed to estimate the temporal and spatial covariance matrices. For each specific signal-to-noise ratio (SNR), 200 Monte Carlo trials are carried out. For comparison, three algorithms are conducted, including the proposed one, and the algorithms in [1,6]. The comparison of the root-mean-square-errors (RMSEs) of the frequency, and elevation and azimuth angle estimates based on the aforementioned algorithms is shown in Figs. 3–5, respectively, for the first scenario, and in Figs. 6–8, respectively, for the second scenario, where the Cramer–Rao lower bound (CRLB) is also provided for reference.

We can note from Figs. 3–8 that the proposed algorithm in general produces close, and sometimes even better, performance in all of the frequency, and elevation and azimuth angle estimates compared with the algorithms in [1,6]. This is due to the fact that although the proposed algorithm only involves the 1-D Unitary ESPRIT, the degradation is compensated by the ingenious combination of estimation and beamforming/filtering processes which sequentially alleviate the contaminated noise as well as refine the estimates stage by stage. Also, the proposed algorithm is inferior to the others in low SNRs because the imprecise estimates in the early stages may affect the accuracy of the succeeding partitioning and of the resulting estimates. However, as we can observe from Figs. 3–8 that this setback is less pronounced in the mid- to high SNRs as now the rough estimates become more accurate.

As for complexity, based on the analytic expressions determined in Remark 2 of Section 3, the number of real multiplications required by the proposed algorithm [1,6] are about 33,177,600, 128,397,312, and 513,598,248, respectively, for the first scenario, and 41,472,000, 128,397,312, and 513,598,248, respectively, for the second scenario. Consequently, the new algorithm requires a substantially lower computational burden.

5. Conclusions

This paper addresses a low-complexity, yet high accuracy algorithm for joint 2-D DOA and frequency estimation. The new algorithm is based on the HSTD technique, which makes full use of the space–time characteristics of the signals by decomposing the signals into finer groups stage by stage in the parameter estimation process. In addition, the data matrices in each stage are appropriately partitioned and restacked so that the new algorithm only calls for a sequence of 1-D Unitary ESPRIT and some temporal filtering/spatial beamforming processes. The incoming signals are thus grouped and estimated, and the pairing of the

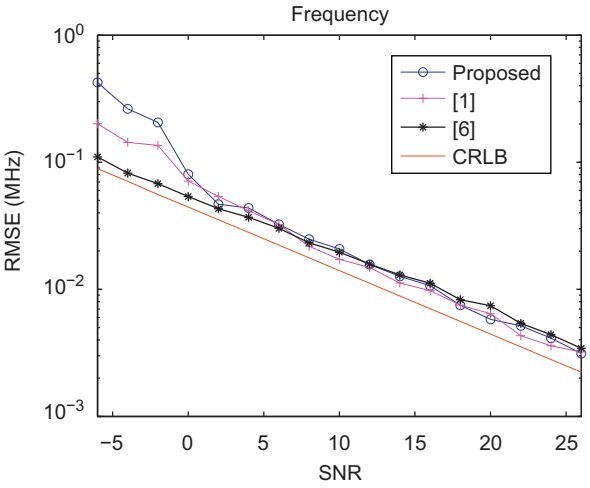

Fig. 3. Comparison of the frequency estimates in the first scenario.

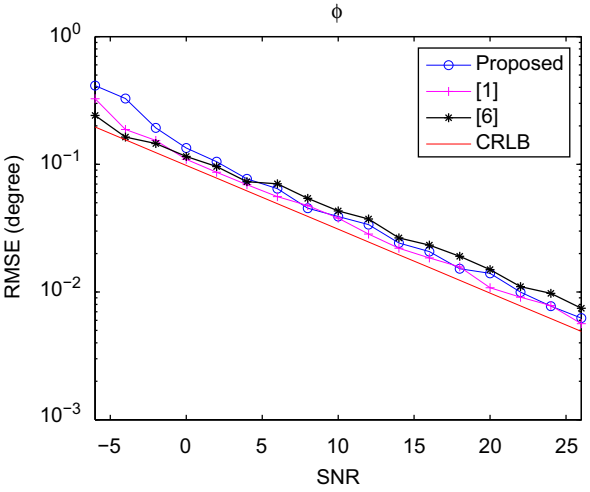

Fig. 4. Comparison of the elevation estimates in the first scenario.

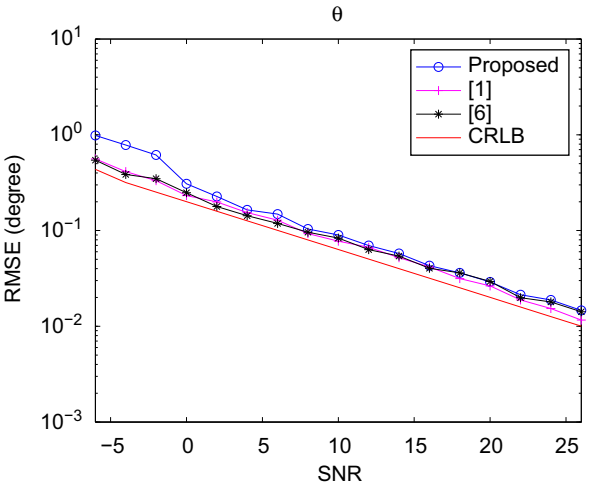

Fig. 5. Comparison of the azimuth estimates in the first scenario.

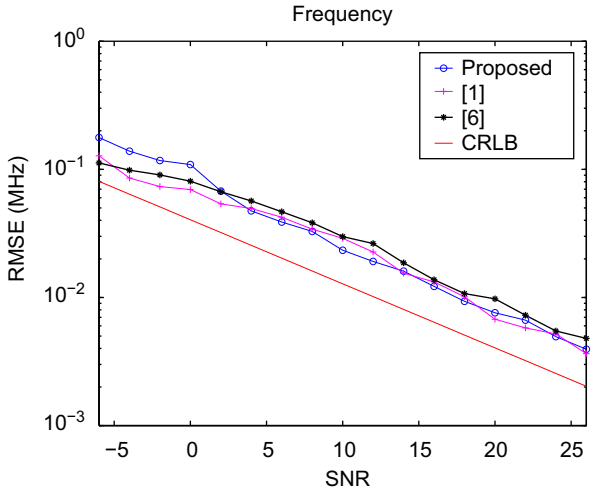

Fig. 6. Comparison of the frequency estimates in the second scenario.

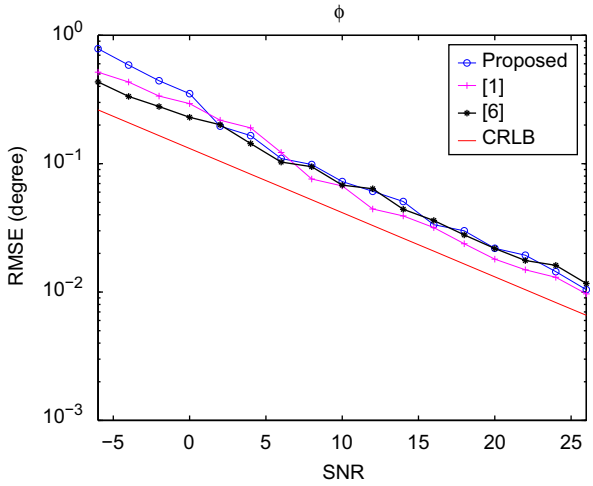

Fig. 7. Comparison of the elevation estimates in the second scenario.

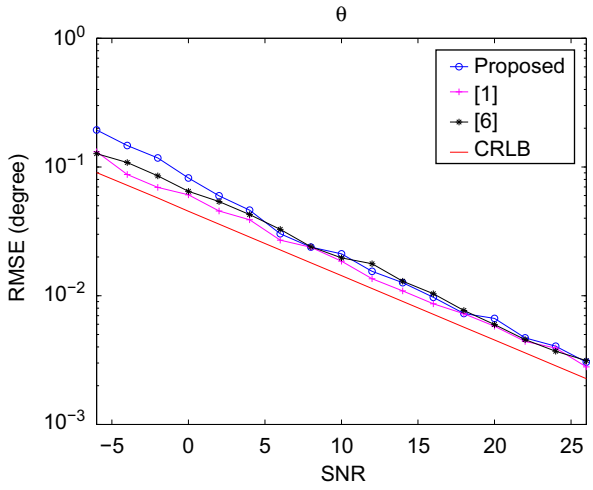

Fig. 8. Comparison of the azimuth estimates in the second scenario.

estimated 2-D DOAs and frequencies is automatically achieved. Conducted simulations show that the new algorithm provides comparable performance as previous works, but with substantially reduced computational overhead.

Acknowledgements

The authors would like to thank the anonymous reviewers for many useful comments which help to improve the quality and readability of this paper.

Appendix A

In this appendix, we will show (13), (19), and (24). We begin with (13). Note that $\mathbf{N}_{u_i}$ can be expressed as

$$\begin{aligned}\mathbf{N}_{u_i} &= [(\mathbf{N}_{f_i}(1:M))^T \vdots \cdots \vdots (\mathbf{N}_{f_i}((N-1)M+1:NM))^T] \\ &= [(\mathbf{P}_{f_i}\mathbf{N}_f(1:M))^T \vdots \cdots \vdots (\mathbf{P}_{f_i}\mathbf{N}_f((N-1)M+1:NM))^T] \quad (29)\end{aligned}$$

Thereby, the (k,p)th element of the noise covariance matrix $\mathbf{R}_{n_{u_i}}$ is given by

$$\begin{aligned}\mathbf{R}_{n_{u_i}}(k,p) &= \frac{1}{LN}\sum_{n=1}^{N}\sum_{l=1}^{L}\mathbf{P}_{f_i}(l)^H E[\mathbf{N}_f((n-1)M+k) \\ &\quad \times(\mathbf{N}_f((n-1)M+p))^H]\mathbf{P}_{f_i}(l) \\ &= \frac{\mathrm{trace}(\mathbf{P}_{f_i})}{L}\sigma^2\delta(k-p), \quad k,p=1,\ldots,M \qquad (30)\end{aligned}$$

where we have used the fact that $\mathbf{P}_{f_i}$ is a projection matrix so $\mathbf{P}_{f_i}^H=\mathbf{P}_{f_i}$ and $(\mathbf{P}_{f_i})^2=\mathbf{P}_{f_i}$, and the white noise assumption, which implies that $E[\mathbf{N}_f((n-1)M+k)(\mathbf{N}_f((n-1)M+p))^H]=\sigma^2\delta(k-p)\mathbf{I}$. The proof is completed by noting that $\mathrm{trace}(\mathbf{P}_{f_i})=\mathrm{rank}(\mathbf{P}_{f_i})=L-q+1$ [13].

Next, we show (19). For this, first note that

$$\begin{aligned}&E[\mathbf{N}_{u_i}((n_1-1)L+k)(\mathbf{N}_{u_i}((n_2-1)L+p))^H] \\ &= E[[\mathbf{P}_{f_i}(k)^H\mathbf{N}_f((n_1-1)M+1)\vdots\cdots\vdots\mathbf{P}_{f_i}(k)^H\mathbf{N}_f(n_1M)]^T \\ &\quad\times[(\mathbf{N}_f((n_2-1)M+1))^H\mathbf{P}_{f_i}(p)\vdots\cdots\vdots(\mathbf{N}_f(n_2M))^H\mathbf{P}_{f_i}(p)]] \\ &= \mathbf{P}_{f_i}(k,p)\sigma^2\delta(n_1-n_2)\mathbf{I}, \quad i=1,\ldots,L, \\ &\quad n_1,n_2=1,\ldots,N, \quad k,p=1,\ldots,L \qquad (31)\end{aligned}$$

where $\mathbf{P}_{f_i}(k,p)=\mathbf{P}_{f_i}(k)^H\mathbf{P}_{f_i}(p)$. Note from (18) that $\mathbf{N}_{v_{ij}}$ can be expressed as

$$\begin{aligned}\mathbf{N}_{v_{ij}} &= [\mathrm{vec}\{(\mathbf{N}_{u_{ij}}(1:L))^T\}\vdots\cdots\vdots\mathrm{vec}\{(\mathbf{N}_{u_{ij}}((N-1)L+1:NL))^T\}]^T \\ &= [\mathrm{vec}\{(\mathbf{P}_{u_{ij}}\mathbf{N}_{u_i}(1:L))^T\} \\ &\quad \vdots\cdots\vdots\mathrm{vec}\{(\mathbf{P}_{u_{ij}}\mathbf{N}_{u_i}((N-1)L+1:NL))^T\}]^T \qquad (32)\end{aligned}$$

Thereby, by using (31), the (k,p)th element of the noise covariance matrix $\mathbf{R}_{n_{v_{ij}}}$ is given by

$$\begin{aligned}\mathbf{R}_{n_{v_{ij}}}(k,p) &= \frac{1}{LM}\sum_{l=1}^{L}\sum_{m=1}^{M}\mathbf{P}_{u_{ij}}(m)^H E[\mathbf{N}_{u_i}((k-1)L+l) \\ &\quad\times(\mathbf{N}_{u_i}((p-1)L+l))^H]\mathbf{P}_{u_{ij}}(m) \\ &= \frac{\mathrm{trace}(\mathbf{P}_{u_{ij}})\mathrm{trace}(\mathbf{P}_{f_i})}{LM}\sigma^2\delta(k-p), \\ &\quad k,p=1,\ldots,N \qquad (33)\end{aligned}$$

The proof of (19) is then established by noting that $\mathrm{trace}(\mathbf{P}_{u_{ij}})=\mathrm{rank}(\mathbf{P}_{u_{ij}})=M-\rho_i+1$ [13]. Therefore, the noise covariance matrices of $\mathbf{N}_{u_i}$ and $\mathbf{N}_{v_{ij}}$ are both diagonal, which implies that the noise components after filtering and restacking are still white.

Finally, we show (24). Again, similar as the above, the outer product of the columns of $\mathbf{N}_{v_{ij}}$ is

$$\begin{aligned}&E[\mathbf{N}_{v_{ij}}((m_1-1)L+k)(\mathbf{N}_{v_{ij}}((m_2-1)L+p))^H] \\ &= E[\mathbf{P}_{u_{ij}}(m_1)^H\mathbf{N}_{u_i}(k)\vdots\cdots\vdots\mathbf{P}_{u_{ij}}(m_1)^H\mathbf{N}_{u_i}((N-1)L+k)]^T \\ &\quad\times[\mathbf{P}_{u_{ij}}(m_2)^H\mathbf{N}_{u_i}(p)\vdots\cdots\vdots\mathbf{P}_{u_{ij}}(m_2)^H\mathbf{N}_{u_i}((N-1)L+p)]^*] \\ &= \mathbf{P}_{f_i}(k,p)\mathbf{P}_{u_{ij}}(m_1,m_2)\sigma^2\mathbf{I} \qquad (34)\end{aligned}$$

for $i=1,\ldots,L$, $j=1,\ldots,\rho_i, m_1$, $m_2=1,\ldots,M,k$, $p=1,\ldots,L$. Based on (23), $\mathbf{N}_{f_{i,j,l}}$ can be re-expressed as

$$\begin{aligned}\mathbf{N}_{f_{i,j,l}} &= [(\mathbf{N}_{v_{i,j,l}}(1:L))^T\vdots\cdots\vdots(\mathbf{N}_{v_{i,j,l}}((M-1)L+1:ML))^T] \\ &= [(\mathbf{P}_{v_{ij,l}}\mathbf{N}_{v_{ij}}(1:L))^T\vdots\cdots\vdots(\mathbf{P}_{v_{ij,l}}\mathbf{N}_{v_{ij,l}}((M-1)L+1:ML))^T] \\ &\qquad (35)\end{aligned}$$

Consequently, by using (34), the (k,p)th element of the noise covariance matrix $\mathbf{R}_{n_{f_{i,j,l}}}$ is given

$$\begin{aligned}\mathbf{R}_{n_{f_{i,j,l}}}(k,p) &= \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\mathbf{P}_{v_{ij,l}}(n)^H E[\mathbf{N}_{v_{ij}}((m-1)L+k) \\ &\quad\times(\mathbf{N}_{v_{ij}}((m-1)L+p))^H]\mathbf{P}_{v_{ij,l}}(n) \\ &= \frac{\mathrm{trace}(\mathbf{P}_{u_{ij}})\mathrm{trace}(\mathbf{P}_{v_{ij,l}})}{MN}\sigma^2\mathbf{P}_{f_i}(k,p), \\ &\quad k,p=1,\ldots,L \qquad (36)\end{aligned}$$

The proof is then established by noting that $\mathrm{trace}(\mathbf{P}_{v_{ij,l}})=\mathrm{rank}(\mathbf{P}_{v_{ij,l}})=N-z_{ij}+1$ [13].

Appendix B

In this appendix, we will show that $\bar{\mathbf{g}}_{i,j,l}=\mathbf{P}_{f_i}\mathbf{g}_{i,j,l}$ in (23) is the eigenvector corresponding to the largest eigenvalue of $\mathbf{R}_{f_{i,j,l}}$. The proof is similar to that in [10]. First, note that post multiplying (25) by $\bar{\mathbf{g}}_{i,j,l}$ yields

$$\begin{aligned}\mathbf{R}_{f_{i,j,l}}\cdot\bar{\mathbf{g}}_{i,j,l} &= (\xi_{i,j,l}^2\bar{\mathbf{g}}_{i,j,l}\bar{\mathbf{g}}_{i,j,l}^H+\sigma_{f_{i,j,l}}^2\mathbf{P}_{f_i})\mathbf{P}_{f_i}\mathbf{g}_{i,j,l} \\ &= (\eta\xi_{i,j,l}^2+\sigma_{f_{i,j,l}}^2)\bar{\mathbf{g}}_{i,j,l} \qquad (37)\end{aligned}$$

where $\eta=\bar{\mathbf{g}}_{i,j,l}^H\bar{\mathbf{g}}_{i,j,l}$ and we have used the fact that $(\mathbf{P}_{f_i})^2=\mathbf{P}_{f_i}$. Eq. (37) implies that $\bar{\mathbf{g}}_{i,j,l}$ is the eigenvector of $\mathbf{R}_{f_{i,j,l}}$ corresponding to the eigenvalue $\eta\xi_{i,j,l}^2+\sigma_{f_{i,j,l}}^2$. Next, we will show that $\eta\xi_{i,j,l}^2+\sigma_{f_{i,j,l}}^2$ is the largest eigenvalue of $\mathbf{R}_{f_{i,j,l}}$. Using the fact that if $\mathbf{A}$ and $\mathbf{B}$ are both $N\times N$ Hermitian matrices, then $\lambda_n(\mathbf{A}+\mathbf{B})\le\lambda_1(\mathbf{A})+\lambda_n(\mathbf{B})$, $n=1,\ldots,N$ [13], where $\lambda_n(\cdot)$ denotes the nth eigenvalue of the matrix inside, and the eigenvalues have been arranged in descending order. Therefore, with $\mathbf{R}_{f_{i,j,l}}$ given in (25), we have

$$\lambda_i(\mathbf{R}_{f_{i,j,l}})\le\lambda_1(\sigma_{i,j}^2\mathbf{P}_{f_{i,j,l}})+\lambda_i(\xi_{i,j,l}^2\bar{\mathbf{g}}_{i,j,l}\bar{\mathbf{g}}_{i,j,l}^H), \quad i=1,\ldots,N \qquad (38)$$

Since both $\mathbf{P}_{f_i}$ and $(1/\eta\xi_{i,j,l}^2)\bar{\mathbf{g}}_{i,j,l}\bar{\mathbf{g}}_{i,j,l}^H$ are Hermitian with eigenvalues $\{\mathbf{1}_{1\times(L-q+1)},\mathbf{0}_{1\times(q-1)}\}$ and $\{1,\mathbf{0}_{1\times(L-1)}\}$, respectively, the largest eigenvalue of $\mathbf{R}_{f_{i,j,l}}$ is then equal to $\eta\xi_{i,j,l}^2+\sigma_{i,j}^2$.

Appendix C

In this appendix, we will prove (27). By definition, the normalized eigenvector $\mathbf{e}_{f_{i,j,l}} = \bar{\mathbf{g}}_{f_{i,j,l}}/\|\bar{\mathbf{g}}_{f_{i,j,l}}\|$ and thus we have

$$\mathbf{e}_{f_{i,j,l}}\mathbf{e}_{f_{i,j,l}}^H = \alpha_{i,j,l}\mathbf{P}_{f_i}\mathbf{g}_{i,j,l}\mathbf{g}_{i,j,l}^H\mathbf{P}_{f_i} \tag{39}$$

where $\alpha_{i,j,l} = 1/\|\bar{\mathbf{g}}_{f_{i,j,l}}\|^2$ and we have used the fact that $\mathbf{P}_{f_i}^H = \mathbf{P}_{f_i}$. Substituting (39) into $\mathbf{\Xi}_{f_{i,j,l}} = \mathbf{I} - (\mathbf{I} - \mathbf{e}_{i,j,l}\mathbf{e}_{i,j,l}^H)\mathbf{P}_{f_i}$, we can obtain

$$\mathbf{\Xi}_{f_{i,j,l}} = \mathbf{I} - \mathbf{P}_{f_i} + \alpha_{i,j,l}\mathbf{P}_{f_i}\mathbf{g}_{i,j,l}\mathbf{g}_{i,j,l}^H\mathbf{P}_{f_i} \tag{40}$$

where we have used (9). On the other hand, using the matrix inversion lemma [12], the projection matrix of $\mathbf{G}_{i,j,l} = [\bar{\mathbf{G}}_i\mathbf{g}_{i,j,l}]$ can be expressed as

$$\begin{aligned}\mathbf{G}_{i,j,l}(\mathbf{G}_{i,j,l}^H\mathbf{G}_{i,j,l})^{-1}\mathbf{G}_{i,j,l}^H &= [\bar{\mathbf{G}}_i\ \ \mathbf{g}_{i,j,l}]\begin{bmatrix}\bar{\mathbf{G}}_i^H\\ \mathbf{g}_{i,j,l}^H\end{bmatrix}[\bar{\mathbf{G}}_i\quad \mathbf{g}_{i,j,l}]^{-1}\begin{bmatrix}\bar{\mathbf{G}}_i^H\\ \mathbf{g}_{i,j,l}^H\end{bmatrix}\\ &= [\bar{\mathbf{G}}_i\ \ \mathbf{g}_{i,j,l}]\begin{bmatrix}\mathbf{C}_{11} & \mathbf{c}_{12}\\ \mathbf{c}_{21}^T & c_{22}\end{bmatrix}\begin{bmatrix}\bar{\mathbf{G}}_i^H\\ \mathbf{g}_{i,j,l}^H\end{bmatrix}\end{aligned} \tag{41}$$

where

$$\begin{aligned}\mathbf{C}_{11} &= (\bar{\mathbf{G}}_i^H\bar{\mathbf{G}}_i)^{-1} + \alpha_{i,j,l}(\bar{\mathbf{G}}_i^H\bar{\mathbf{G}}_i)^{-1}\bar{\mathbf{G}}_i^H\mathbf{g}_{i,j,l}\mathbf{g}_{i,j,l}^H\bar{\mathbf{G}}_i(\bar{\mathbf{G}}_i^H\bar{\mathbf{G}}_i)^{-1}\\ \mathbf{c}_{12} &= -\alpha_{i,j,l}(\bar{\mathbf{G}}_i^H\bar{\mathbf{G}}_i)^{-1}\bar{\mathbf{G}}_i^H\mathbf{g}_{i,j,l}\\ \mathbf{c}_{21}^T &= \mathbf{c}_{12}^H\\ c_{22} &= \alpha_{i,j,l}\end{aligned} \tag{42}$$

By substituting (42) into (41) and after some manipulations, Eq. (41) can be readily shown as

$$\begin{aligned}\mathbf{G}_{i,j,l}(\mathbf{G}_{i,j,l}^H\mathbf{G}_{i,j,l})^{-1}\mathbf{G}_{i,j,l}^H &= \mathbf{I} - \mathbf{P}_{f_i} + \alpha_{i,j,l}\mathbf{g}_{i,j,l}\mathbf{g}_{i,j,l}^H\\ &\quad - \alpha_{i,j,l}(\mathbf{I} - \mathbf{P}_{f_i})\mathbf{g}_{i,j,l}\mathbf{g}_{i,j,l}^H\\ &\quad - \alpha_{i,j,l}\mathbf{g}_{i,j,l}\mathbf{g}_{i,j,l}^H(\mathbf{I} - \mathbf{P}_{f_i})\\ &\quad + \alpha_{i,j,l}(\mathbf{I} - \mathbf{P}_{f_i})\mathbf{g}_{i,j,l}\mathbf{g}_{i,j,l}^H(\mathbf{I} - \mathbf{P}_{f_i})\\ &= \mathbf{I} - \mathbf{P}_{f_i} + \alpha_{i,j,l}\mathbf{P}_{f_i}\mathbf{g}_{i,j,l}\mathbf{g}_{i,j,l}^H\mathbf{P}_{f_i}\end{aligned} \tag{43}$$

which is equal to (40) and thus completes the proof.

References

[1] M. Haardt, J.A. Nossek, 3-D unitary ESPRIT for joint 2-D angle and carrier estimation, in: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 1997, pp. 255–258.

[2] L.C. Godara, Application of antenna arrays to mobile communications, part II: beam-forming and direction-of-arrival considerations, Proc. IEEE 85 (8) (August 1997) 1195–1245.

[3] R.D. Balakrishnan, B.S. Nugroho, H.M. Kwon, Joint DOA estimation and multi-user detection for WCDMA using DELSA, in: Proceedings of the IEEE International Symposium on Spread Spectrum Techniques and Applications, 2004, pp. 879–884.

[4] R. Roy, T. Kailath, ESPRIT—estimation of signal parameters via rotational invariance techniques, IEEE Trans. Acoust. Speech. Signal Process. 37 (7) (July 1989) 984–995.

[5] M. Haardt, J.A. Nossek, Unitary ESPRIT: how to obtain increased estimation accuracy with a reduced computational burden, IEEE Trans. Signal Process. 43 (5) (May 1995) 1232–1242.

[6] P. Strobach, Total least squares phased averaging and 3-D ESPRIT for joint azimuth-elevation-carrier estimation, IEEE Trans. Signal Process. 49 (1) (January 2001) 54–62.

[7] C.-Y. Qi, Y.-L. Wang, Y.-S. Zhang, H. Chen, A space–time processing approach to joint estimation of signal frequency and 2-D direction of arrival, in: Proceedings of the IEEE AP-S/URSI, vol. 3A, 2005, pp. 370–373.

[8] N. Tayem, H.M. Kwon, S. Min, D. Kang, FOA and 2-D DOA estimation with propagator method, in: Proceedings of the IEEE International Conference on Vehicular Technology Conference, vol. 4, 2005, pp. 2211–2215.

[9] Y.-Y. Wang, J.-T. Chen, W.-H. Fang, TST—MUSIC for joint DOA-delay estimation, IEEE Trans. Signal Process. 46 (4) (April 2001) 721–729.

[10] J.-D. Lin, W.-H. Fang, Y.-Y. Wang, J.-T. Chen, FSF MUSIC for joint DOA and frequency estimation and its performance analysis, IEEE Trans. Signal Process. 54 (12) (December 2006) 4529–4542.

[11] C.-H. Lin, W.-H. Fang, K.-H. Wu, J.-D. Lin, Fast algorithm for joint azimuth and elevation angles, and frequency estimation via hierarchical space–time decomposition, in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Hawaii, USA, 2007, pp. 1061–1064.

[12] H.L. Van Trees, Optimum Array Processing, Wiley-Interscience, 2002.

[13] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, 1985.

[14] A.N. Lemma, A.-J. van der Veen, E.F. Deprettere, Analysis of joint angle-frequency estimation using ESPRIT, IEEE Trans. Signal Process. 51 (5) (May 2003) 1264–1283.

[15] M. Wax, T. Kailath, Detection of signals by information theoretic criteria, IEEE Trans. Acoust. Speech Signal Process. 33 (2) (April 1985) 387–392.