Joint DOA and frequency estimation with sub-Nyquist sampling

Liang Liu [a,b,*], Jian-Feng Gu [c], Ping Wei [b]

[a] Southwest China Institute of Electronic Technology, Chengdu 610036, China
[b] School of Information Communication Engineering, Center for Cyber Security, University of Electronic Science and Technology of China, Chengdu 611731, China
[c] Faculty of Electrical Engineering and ComputerScience, Ningbo University, 315211, China, and Mistronix, 2500 Chemin de Polytechnique, Montréal, QC H3T1J4, Canada

ABSTRACT

Inspired by the concept of sub-Nyquist sampling, in this paper, a receiver architecture of sensor arrays for jointly estimating the frequency and direction-of-arrival (DOA) of the narrowband far-field signals is presented. In particular, our contribution is threefold. (i) First, we propose a time-space unionized signal reception model for array signals, where sub-Nyquist sampling techniques can be used. (ii) Second, based on the reception model, two algorithms for jointly estimating frequency and DOA are proposed. One is based on a trilinear decomposition by the third-order tensor theory and the other is based on subspace decomposition. (iii) Third, we derive the corresponding Cramér-Rao Bound (CRB) for such a model and analyze the computational complexity of the proposed algorithms as well. Simulation results demonstrate that the estimation performance of our proposed algorithms based on the receiver architecture closely match the CRBs with respect to the noise levels, the branch number and the source number.

© 2018 Elsevier B.V. All rights reserved.

1. Introduction

With the rapid development of the number of users and their data rates, the traditional fixed spectrum policy is no longer a feasible approach. One of potential techniques for alleviating the spectrum scarcity is the use of spectrum sensing to detect the unused spectrum holes over some frequency ranges by using Cognitive Radio (CR) technology [1–7]. In addition, multiple antenna techniques are widely used in communications, e.g., massive MIMO for 5G wireless communications [8–10], and their effectiveness have been shown in different aspects [11–19]. In this way, the direction of arrival (DOA) can be thought as an additional spatial spectrum to improve the spectrum sensing performance in CRs. Therefore, multiple antenna spectrum sensing techniques have been considered in [20–22], where the carrier frequencies associated with DOAs of multiple signals were estimated simultaneously. Unfortunately, there are at least two shortcomings. One is the pair matching problem for the carrier frequencies and the corresponding DOAs. The other is that time domain sampling rate is equal to or larger than the Nyquist sampling rate which is considered as a bottleneck for wideband signal processing especially in CRs. For instance, it leads to prohibitive Nyquist sampling rate and massive sampling

data to be processed if the spectrum needed to be monitored from 300 MHz to several GHz [1,3,4,6,7].

To deal with the problem of high sampling rate, the compressive sampling (CS) techniques have been proposed and applied in many practical applications [23–28]. In particular, there is an extensive literature on such techniques in the array signal processing area [29–37]. Gu et al. [29–32] carried out the beamforming, DOA estimation, and time-delay estimation based on the CS techniques, respectively. Inspired by the concept of sub-Nyquist sampling, the authors presented the methods to jointly estimate the carrier frequency and DOA of multiple signals in [33–37]. A multiband signal model is reconstructed on the data obtained below the Nyquist sampling rate [25,26,38,39]. Ariananda and Leus [33] suggested a new structure, where each output of a linear array is carried out through the multi-coset sampling. Later, the minimum redundancy array (MRA) is employed to estimate the DOA of more uncorrelated sources than sensors in [33]. The wide-sense stationary signal can be compressed in both the time domain and the spatial domain. However, the accuracy of frequency and DOA estimate is constrained by the inverse of block length and array aperture, respectively, and a peak-searching technique for two-dimensional (2-D) parameters' space is needed to get the frequency and DOA estimates.

To simplify the complexity of hardware implementations, AnilKumar et al. [34] added an identical delayed channel for each sensor. Unfortunately, the problem of pair matching happened for

---

* Corresponding author at: Southwest China Institute of Electronic Technology, Chengdu 610036, China.
E-mail address: liu_yinliang@outlook.com (L. Liu).

the use of the uniform linear array (ULA). To solve such a problem, they also proposed another structure in [35] enjoying the same hardware complexity. In [36], the authors proposed the so-called space-time array, each sensor of which is connected by the same multiple delayed channels. In this scheme, simultaneous frequency and DOA estimation can be achieved in the case of more sources than sensors. The estimation performance of [34–36] is not good enough and the array configurations of [35,36] are also constrained by the ULA system. More recently, Ioushua et al. [37] proposed an alternative sub-Nyquist sampling and signal reconstruction method by the modulated wideband converter (MWC) for ULA. Based on the scheme, they also developed a method for estimating both the operating frequency and the DOA via an L-shaped array.

The aim of this paper is to jointly estimate the carrier frequency and DOA of multiple narrowband far-field signals, where the multi-coset sampling [26] is applied to each sensor of the array. It is worth noting that arbitrary arrays are set up to gain certain advantage, for instance, the MRA can achieve a higher estimation accuracy than the ULA does with the same sensor number. To improve the performance of joint estimation on the carrier frequency and DOA, we elaborately construct a novel time-space signal model. By using such a model we can extend either the temporal space or the spatial space to a whole time-space unionized space. In addition, the Cramér-Rao Bound (CRB) for spatial phase estimation is also derived for this model. It is proven that the CRB is not affected by the number of signals provided that the average sampling rate in our architecture is consistent with Nyquist rate, while the CRB based on Nyquist sampling increases in terms of the number of signals. Finally, the simulations verify the above conclusions on the CRB and the superior performance of the proposed methods with respect to the noise levels, the number of branches, and the number of source signals as well.

This paper is organized as follows. In the following section, we present the fundamental array signal model and the assumptions. In Section 3, a receiver architecture is presented and the corresponding signal reception model is derived. Based on such a signal model, two joint carrier frequency and DOA estimation methods are proposed in Section 4. In Section 5, the CRB is derived and analyzed. The computer simulations are carried out in Section 6. Section 7, the conclusions of our findings are given.

The following notations are used in the paper. $(\cdot)^{\mathrm{T}}$, $(\cdot)^{\mathrm{H}}$, and $(\cdot)^{\dagger}$ denote the transpose, Hermitian transpose, and Moore–Penrose pseudo-inverse, respectively. $\mathrm{E}(\cdot)$ stands for the expectation operator. x_j is the jth entry of a vector $\mathbf{x}$. $\mathbf{A}_i$, $\mathbf{A}^j$, and A_{ij} are the ith row, the jth column, and (i, j)th entry of a matrix $\mathbf{A}$, respectively. $\otimes$, $\odot$, and $*$ denote the Kronecker product, Hadamard product, and Khatri-Rao product, respectively. $\mathbf{I}_M$ stands for an $M \times M$ identity matrix.

2. Array signal model and fundamental assumptions

Consider K narrowband far-field signals modeled as multiband signals [26] impinging on an array composed of M ($M > K$) sensors, where the operating frequency of each signal is located widely. The array output can be written as [11]

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t), \tag{1}$$

where $\mathbf{s}(t) = [s_1(t), \ldots, s_K(t)]^{\mathrm{T}}$ is the uncorrelated signal vector and $\mathbf{n}(t) = [n_1(t), \ldots, n_M(t)]^{\mathrm{T}}$ is assumed to be a i.i.d complex Gaussian noise with zero-mean and covariance matrix of $\sigma^2 \mathbf{I}_M$. The general linear array manifold matrix has the form as $\mathbf{A} = [\mathbf{a}(\phi_1), \ldots, \mathbf{a}(\phi_K)]$. Note that an ambiguity problem maybe occur for linear arrays. Readers can find some details in [40–42] about this problem, which can be handled by the sample covariance matrix augmenting techniques [43,44] or the idea suggested in [45]. For simplicity, we assume herein that
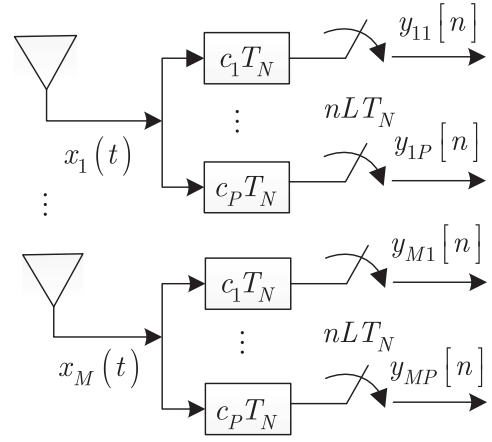

Fig. 1. Proposed receiver architecture.

$\mathbf{A}$ is a full rank matrix without the problem. Without loss of generality, the steering vector of an linear array is $\mathbf{a}(\phi_k) = [\exp(j\phi_k d_1), \ldots, \exp(j\phi_k d_M)]^{\mathrm{T}}$ with the sensor positions of $\mathbf{d} = [d_1, d_2, \ldots, d_M]d$, where the spatial phase

$$\phi_k = \frac{2\pi d \sin(\theta_k)}{\lambda_k} = \frac{\pi \sin(\theta_k) f_k}{f_N}. \tag{2}$$

d is the half-wavelength in terms of the Nyquist sampling rate f_N. θ_k and f_k are the DOA and the carrier frequency of $s_k(t)$, respectively. Thus, the frequency domain output can be expressed as

$$\mathbf{X}(f) = \mathbf{A}\mathbf{S}(f) + \mathbf{N}(f), \tag{3}$$

where $\mathbf{X}(f) = [X_1(f), \ldots, X_M(f)]^{\mathrm{T}}$, $\mathbf{S}(f) = [S_1(f), \ldots, S_K(f)]^{\mathrm{T}}$, and $\mathbf{N}(f) = [N_1(f), \ldots, N_M(f)]^{\mathrm{T}}$ are the frequency domain expression of $\mathbf{x}(t)$, $\mathbf{s}(t)$, $\mathbf{n}(t)$, respectively.

Remark 1. The basic assumptions herein are presented to introduce the proposed algorithms clearly since the sub-Nyquist sampling is induced. (i) Narrowband far-field signals are used, i.e., the signal bandwidth much less than $f_{\mathrm{sub}} = f_N/L$, where L is the sampling rate reduction factor of sub-Nyquist sampling and f_N is the Nyquist sampling rate. (ii) There is at most one signal in each subband if the whole frequency band is divided into L subbands of equal length, which often happened in the case of widely separated signals. (iii) The number of signals K is known *apriori* or is estimated by related methods, such as AIC or MDL methods. The objective of this paper is to estimate the carrier frequency f_k and DOA θ_k of multiple signals $s_k(t)$ simultaneously.

3. Proposed receiver architecture and signal reception model

In this section, the architecture is presented and the corresponding signal reception model is derived.

3.1. Receiver architecture

Similar to the multi-coset sampling suggested in [26], each sensor of the array with M elements is connected by P delay branches shown in Fig. 1. The whole analog-to-digital converters (ADCs) are well-synchronized with a sub-Nyquist sampling rate, namely $f_{\mathrm{sub}} = f_N/L$, where L is so-called the undersampling factor. The sampling pattern belongs to the integer set of $C = [c_1, \ldots, c_P]$ with $0 \le c_1 < c_2 < \cdots < c_P \le L - 1$. The average sampling rate of the multi-coset sampling is

$$f_E = \frac{P f_N}{L}, \tag{4}$$

which is lower than the Nyquist rate f_N when $P < L$.

3.2. Signal reception model

Recalling the results in [26], we can obtain the following relationship

$$Y_{mp}\left(e^{j2\pi f T_N}\right) = \frac{1}{LT_N} \sum_{l=1}^{L} \exp\left(j\frac{2\pi}{L} c_p(l-1)\right) X_{ml}(f),$$

$$1 \le p \le P, f \in \mathcal{F} \triangleq [0, f_{\text{sub}}), \tag{5}$$

where $Y_{mp}\left(e^{j2\pi f T_N}\right)$ and $X_m(f)$ are the Fourier transform of $y_{mp}[n]$ and $x_m(t)$. The matrix form of (5) is expressed as

$$\mathbf{Y}_m(f) = \mathbf{B}\overline{\mathbf{X}}_m(f), f \in \mathcal{F}, 1 \le m \le M, \tag{6}$$

where $\overline{\mathbf{X}}_m(f) = [X_{m1}(f), \ldots, X_{mL}(f)]^T$, $X_{ml}(f) = X_m(f + (l-1)f_{\text{sub}})$, $B_{il} = \frac{1}{LT_N} \exp\left(j\frac{2\pi}{L} c_i(l-1)\right)$, $\mathbf{Y}_m(f) = \left[Y_{m1}\left(e^{j2\pi f T_N}\right), \ldots, Y_{mP}\left(e^{j2\pi f T_N}\right)\right]^T$. To simplify the expressions, we can redefine $B_{il} \triangleq \frac{1}{\sqrt{L}} \exp\left(j\frac{2\pi}{L} c_i(l-1)\right)$ and $Y_{mp}(f) \triangleq \sqrt{L}T_N Y_{mp}\left(e^{j2\pi f T_N}\right)$ by multiplying $\sqrt{L}T_N$ on both sides of (6). The ith row of $\mathbf{B}$ can be expressed as

$$\mathbf{B}_i = \frac{1}{\sqrt{L}}\left[\exp\left(j\frac{2\pi}{L} c_i 0\right), \cdots, \exp\left(j\frac{2\pi}{L} c_i(L-1)\right)\right].$$

Therefore, it is easy to verify that

$$\mathbf{B}_i \mathbf{B}_j^H = \frac{1}{L} \sum_{l=1}^{L} \exp\left(j\frac{2\pi}{L}\left(c_i - c_j\right)(l-1)\right)$$

$$= \begin{cases} 1, & i = j, \\ 0, & i \ne j. \end{cases} \tag{7}$$

(7) implies that $\mathbf{B}$ is row orthogonal, i.e.,

$$\mathbf{B}\mathbf{B}^H = \mathbf{I}_P. \tag{8}$$

Mishali and Eldar [26] show us the sufficient condition for an unique solution in (6): The sampling pattern C is an entire pattern with a fully Kruskal-rank $\mathbf{B}$ and parameters P and L satisfy $P \ge K$ and $L \le 1/BT_N$, where B is the maximum signal bandwidth.

From (3),

$$X_{ml}(f) = \mathbf{A}_m \widehat{\mathbf{S}}_l(f) + N_{ml}(f), 1 \le l \le L, f \in \mathcal{F}, \tag{9}$$

where $\widehat{\mathbf{S}}_l(f) = [S_{1l}(f), \ldots, S_{Kl}(f)]^T$, $S_{kl}(f) = S_k(f + (l-1)f_{\text{sub}})$, $N_{ml}(f) = N_m(f + (l-1)f_{\text{sub}})$. In matrix form, (9) can be expressed as

$$\overline{\mathbf{X}}_m(f) = (\mathbf{I}_L \otimes \mathbf{A}_m)\widehat{\mathbf{S}}(f) + \widehat{\mathbf{N}}_m(f), f \in \mathcal{F}, \tag{10}$$

where $\widehat{\mathbf{S}}(f) = \left[\widehat{\mathbf{S}}_1^T(f), \ldots, \widehat{\mathbf{S}}_L^T(f)\right]^T \in \mathbb{C}^{KL \times 1}$, $\widehat{\mathbf{N}}_m(f) = [N_{m1}(f), \ldots, N_{mL}(f)]^T$. Substituting (10) into (6), we get

$$\mathbf{Y}_m(f) = \mathbf{B}(\mathbf{I}_L \otimes \mathbf{A}_m)\widehat{\mathbf{S}}(f) + \mathbf{B}\widehat{\mathbf{N}}_m(f)$$

$$= (\mathbf{A}_m \otimes \mathbf{B})\overline{\mathbf{S}}(f) + \mathbf{B}\widehat{\mathbf{N}}_m(f), f \in \mathcal{F}, 1 \le m \le M, \tag{11}$$

where $\overline{\mathbf{S}}(f) = \left[\overline{\mathbf{S}}_1^T(f), \ldots, \overline{\mathbf{S}}_K^T(f)\right]^T \in \mathbb{C}^{KL \times 1}$, $\overline{\mathbf{S}}_k(f) = [S_{k1}(f), \ldots, S_{kL}(f)]^T$. Then, combining all m can result in

$$\mathbf{Y}(f) = (\mathbf{A} \otimes \mathbf{B})\overline{\mathbf{S}}(f) + (\mathbf{I}_M \otimes \mathbf{B})\widehat{\mathbf{N}}(f) \tag{12}$$

$$\triangleq \mathbf{G}\overline{\mathbf{S}}(f) + \mathbf{I}_{\mathbf{B}}\widehat{\mathbf{N}}(f), f \in \mathcal{F}, \tag{13}$$

where $\mathbf{G} \triangleq \mathbf{A} \otimes \mathbf{B} \in \mathbb{C}^{MP \times KL}$, $\mathbf{I}_{\mathbf{B}} \triangleq \mathbf{I}_M \otimes \mathbf{B} \in \mathbb{C}^{MP \times ML}$, $\mathbf{Y}(f) = [\mathbf{Y}_1^T(f), \ldots, \mathbf{Y}_M^T(f)]^T \in \mathbb{C}^{MP \times 1}$, $\widehat{\mathbf{N}}(f) = [\widehat{\mathbf{N}}_1^T(f), \ldots, \widehat{\mathbf{N}}_M^T(f)]^T \in \mathbb{C}^{ML \times 1}$.

Since the carrier frequencies are distributed in the different sub-bands and $s_k(t)$ is assumed as a narrowband signal, i.e., l_i belonging to the activated index set of $\Omega = [l_1, \ldots, l_K]$ is not equal to l_j for any $i \ne j$, $\overline{\mathbf{S}}(f)$ is a K-sparse vector of length KL. The support index $\mathcal{S}$ of $\overline{\mathbf{S}}(f)$ and $\mathbf{G}$ is determined as

$$\mathcal{S}_k = (k-1)L + l_k. \tag{14}$$

Thus, (13) can be written as

$$\mathbf{Y}(f) = \mathbf{G}_{\mathcal{S}}\overline{\mathbf{S}}^{\mathcal{S}}(f) + \mathbf{I}_{\mathbf{B}}\widehat{\mathbf{N}}(f) \tag{15}$$

$$= (\mathbf{A} * \mathbf{B}_{\Omega})\overline{\mathbf{S}}^{\mathcal{S}}(f) + \mathbf{I}_{\mathbf{B}}\widehat{\mathbf{N}}(f), f \in \mathcal{F}. \tag{16}$$

where $\mathbf{G}_{\mathcal{S}} \triangleq \mathbf{A} * \mathbf{B}_{\Omega} \in \mathbb{C}^{MP \times K}$, $\mathbf{B}_{\Omega} \in \mathbb{C}^{P \times K}$, $\overline{\mathbf{S}}^{\mathcal{S}}(f) \in \mathbb{C}^{K \times 1}$.

In (15), since the Khatri-Rao product is used to combine the frequency domain and spatial domain into a 2-D matrix form, this equation can be solved by a third-order tensor problem. On the other hand, if $\mathbf{G}_{\mathcal{S}}$ is constructed from a special array manifold, the subspace decomposition technique can be implemented. Therefore, we derive two different methods on the different point of views as well analyze the performance in detail.

Remark 2. It will result in the pair-matching problem and poor target classification performance if the frequency and DOA are estimated individually on the sub-Nyquist sampling model (6) and DOA model (9). Conversely, the target classification can be effectively performed in a time-space unionized reception model (15) without the pair-matching problem, where the compressive sampling technique can be applied in time-space domain, respectively.

Remark 3. The model of [36] is constructed only in the time domain, where the temporal manifold matrix in [36] can be considered as the spatial steering vectors in array signal processing literatures. The frequency information is involved in each column of temporal manifold matrix. However, the model (15) is formed in frequency domain, which separates the different frequency band by the multi-coset sampling concept. In this way, the sampling of each time-delay channel contains a mixed observation of the original signal associated with the coefficient determined by the compressed matrix $\mathbf{B}$, which is the key point of compressed sampling. The different modeling schemes result in the different constraints of the parameters. In other words, the consecutive delay channels are needed in [36] while the sampling pattern should satisfy the constraint of the fully Kruskal-rank of the corresponding matrix $\mathbf{B}$ in our model. In [34,35], the data models of delay channels and array channels are constructed individually rather than jointly. Besides, the techniques of [35,36] are deduced in the ULA system, where the spatial compressive sampling can not be implemented. In [33], a prohibitive 2-D searching technique is needed to obtain the accurate frequency and DOA estimates.

4. Joint DOA and frequency estimation algorithm

4.1. Algorithm based on trilinear decomposition

It is easy to express (15) in an element-wise form as

$$Y_{mp}(f) = \sum_{k=1}^{K} A_{mk} B_{pl_k} S_{kl_k}(f), f \in \mathcal{F}. \tag{17}$$

From (17), $Y_{mp}(f)$ can be considered as a third-order tensor which can be solved by trilinear decomposition technique [46,47]. Although the problem is different from the standard trilinear decomposition problem due to known $\mathbf{B}$, we can still use the standard trilinear decomposition algorithm, such as alternating least squares (ALS) [48] and regularized alternating least squares (RALS) [49,50] to obtain $\widehat{\mathbf{A}}$, $\widetilde{\mathbf{B}}$, and $\widetilde{\mathbf{S}}(f)$, which are the estimates of $\mathbf{A}$, $\mathbf{B}_{\Omega}$, and $\overline{\mathbf{S}}^{\mathcal{S}}(f)$, respectively.

Table 1
Algorithm JDFTD.

Step	Operation
(a)	Obtain $\widetilde{\mathbf{A}}$, $\widetilde{\mathbf{B}}$, and $\widetilde{\mathbf{S}}(f)$ using ALS / RALS according to (17);
(b)	Gain ϕ_k applying (18) to $\widetilde{\mathbf{A}}$;
(c)	Determine Ω_k according to (21);
(d)	Get $\widetilde{\mathbf{s}}_k(t)$ and $\overline{f}_k$ by applying IFFT, periodogram-like technique to $\widetilde{\mathbf{S}}_k(f)$, $\widetilde{\mathbf{s}}_k(t)$, consecutively;
(e)	Compute f_k through (22);
(f)	Calculate θ_k through (2);

Herein, the trilinear decomposition of model (17) is identifiable except for permutation and scaling, which are robust to our algorithms. The following facts should be used: (i) $\mathbf{A}$ is a full column rank matrix, so $k_{\mathbf{A}} = K$, where $k_{\mathbf{A}}$ is the Kruskal-rank of $\mathbf{A}$ [46]. (ii) $\mathbf{B}_{\Omega}$ is a full column rank matrix and $k_{\mathbf{B}_{\Omega}} = K$ because $\mathbf{B}$ is a fully Kruskal-rank matrix, i.e., $k_{\mathbf{B}} = P$, and $|\widetilde{\Omega}|=K \leq P$. (iii) Since $\mathbf{Q} = \left(\overline{\mathbf{S}}^S\right)^T$, which is the matrix form of $\overline{\mathbf{S}}^S(f)$ in the same block, i.e., $f \in \mathcal{F}$, the signals have been supposed to be uncorrelated to make sure full column rank of $\mathbf{Q}$. Furthermore, we hold $k_{\mathbf{Q}} = K$. Combining these facts, we obtain $k_{\mathbf{A}}+k_{\mathbf{B}_{\Omega}} + k_{\mathbf{Q}}=3K$. We thus have $k_{\mathbf{A}}+k_{\mathbf{B}_{\Omega}} + k_{\mathbf{Q}}=3$ when $K=1$, and $k_{\mathbf{A}}+k_{\mathbf{B}_{\Omega}} + k_{\mathbf{Q}} \geq 2K+2$ when $K \geq 2$. The trilinear decomposition of model (17) is identifiable according to Theorem 4i and Theorem 4a in [46], respectively.

Since every column vector of $\widetilde{\mathbf{A}}$ contains one pair of the carrier frequencies and DOAs, the maximum likelihood (ML) estimation for ϕ_k is processed by

$$\hat{\phi}_k = \arg \max_{\phi \in [0,2\pi)} \left|\mathbf{a}^{\mathrm{H}}(\phi)\widetilde{\mathbf{A}}^k\right|. \qquad (18)$$

According to the above discussions, we can observe that the trilinear decomposition of model (17) is identifiable except that the order and scale of the columns of the estimated matrices are ambiguous. Since $\mathbf{B}$ is a fully Kruskal-rank matrix and $\widetilde{\mathbf{B}}$ is a full column rank matrix. There is one and only one column in $\mathbf{B}$ which is linearly related to the each column of the matrix $\widetilde{\mathbf{B}}$ in the absence of noise, i.e., the absolute value of the correlation coefficient between them is 1 and it is the potential maximum value. Let us take an example as follows

$$\mathbf{B} = \frac{1}{\sqrt{3}}\begin{bmatrix} \exp\left(j\frac{2\pi}{3}0\right) & \exp\left(j\frac{2\pi}{3}0\right) & \exp\left(j\frac{2\pi}{3}0\right) \\ \exp\left(j\frac{2\pi}{3}0\right) & \exp\left(j\frac{2\pi}{3}1\right) & \exp\left(j\frac{2\pi}{3}2\right) \end{bmatrix} \qquad (19)$$

and

$$\widetilde{\mathbf{B}} = \begin{bmatrix} \exp\left(j\frac{2\pi}{3}0\right) & \exp\left(j\frac{2\pi}{3}0\right) \\ \exp\left(j\frac{2\pi}{3}2\right) & \exp\left(j\frac{2\pi}{3}0\right) \end{bmatrix}. \qquad (20)$$

Only the third and first columns of $\mathbf{B}$ are linearly related to the first and second columns of the matrix $\widetilde{\mathbf{B}}$, respectively. Therefore, we determine Ω by comparing the correlation coefficient of the column $\mathbf{B}$ and $\widetilde{\mathbf{B}}$ as

$$\Omega_k = \arg \max_j r_{kj} = \left|\frac{\left(\widetilde{\mathbf{B}}^k\right)^{\mathrm{H}}\mathbf{B}^j}{\left\|\widetilde{\mathbf{B}}^k\right\|\left\|\mathbf{B}^j\right\|}\right|, j = 1, \cdots, L. \qquad (21)$$

Similarly, the frequency $\overline{f}_k$ of the overlapped signal $\widetilde{\mathbf{s}}_k(t)$ can be estimated by applying the periodogram-like technique to the overlapped signal, and then the original signal's frequency f_k can be obtained by the relationship

$$f_k = (\Omega_k - 1)\frac{f_N}{L} + \overline{f}_k. \qquad (22)$$

Finally, θ_k can be acquired by (2) after ϕ_k and f_k are estimated. Herein, such a joint algorithm based on trilinear decomposition (JDFTD) is outlined in Table 1. It is well known that the use of the trilinear decomposition in (17) leads to an ambiguous problem in terms of permutation and scaling, fortunately, the proposed method can estimate the parameters without ambiguity because those parameters are related to the phase and subband information.

4.2. Algorithm based on subspace decomposition

Making use of (8), we have $\mathbf{I}_B\mathbf{I}_B^{\mathrm{H}} = (\mathbf{I}_M \otimes \mathbf{B})(\mathbf{I}_M \otimes \mathbf{B})^{\mathrm{H}} = \mathbf{I}_M \otimes \left(\mathbf{BB}^{\mathrm{H}}\right) = \mathbf{I}_{MP}$, and $\mathrm{E}\left(\mathbf{I}_B\widehat{\mathbf{N}}(f)\left(\mathbf{I}_B\widehat{\mathbf{N}}(f)\right)^{\mathrm{H}}\right)=\sigma^2\mathbf{I}_{MP}$. Therefore, the noise $\mathbf{I}_B\widehat{\mathbf{N}}(f)$ in (15) is white. The covariance matrix of $\mathbf{Y}(f)$, $f \in \mathcal{F}$ is thus given by

$$\mathbf{R} = \mathrm{E}\left(\mathbf{Y}(f)\mathbf{Y}^{\mathrm{H}}(f)\right)=\mathbf{G}_S\mathbf{R}_{\overline{S}}\mathbf{G}_S^{\mathrm{H}} + \sigma^2\mathbf{I}_{MP}, \qquad (23)$$

where $\mathbf{R}_{\overline{S}}$ is the covariance matrix of the signal sources $\overline{\mathbf{S}}(f)$. The size of the support index $\mathcal{S}$ of $\overline{\mathbf{S}}(f)$ is K, which indicates that the rank of $\mathbf{R}_{\overline{S}}$ is K.

In practical applications the estimated covariance matrix can be given as

$$\mathbf{R} = \frac{1}{T/L}\sum_{f=1}^{T/L}\mathbf{Y}(f)\mathbf{Y}^{\mathrm{H}}(f) \qquad (24)$$

when the snapshots T are large enough.

Implementing the singular value decomposition (SVD) for $\mathbf{R}$ results in

$$\mathbf{R} = \mathbf{U}_S\mathbf{D}_S\mathbf{U}_S^{\mathrm{H}} + \mathbf{U}_N\mathbf{D}_N\mathbf{U}_N^{\mathrm{H}}, \qquad (25)$$

where $\mathbf{U}_S$ and $\mathbf{U}_N$ are the signal subspace and the noise subspace, respectively. Since the signal subspace and the noise subspace are orthogonal, i.e., $\mathbf{a}_l(\phi)\perp\mathbf{U}_N$, where $\mathbf{a}_l(\phi) = \mathbf{a}(\phi) * \mathbf{B}_l$, a 2-D peak-searching algorithm is carried on the following pseudo-spectra

$$P(l, \phi) = \frac{1}{\left\|\mathbf{a}_l^{\mathrm{H}}(\phi)\mathbf{U}_N\right\|^2} \qquad (26)$$

to obtain the estimates of ϕ_k, l_k. It is worth noting that the 2-D searching problem in (26) is different from the traditional 2-D searching problem in [33]. Specifically, the former is to set a fine grid only in the spatial domain because the sub-band position l is a positive integer within $L - 1$, where only one continuous parameter is needed to be discretized. However, the latter one is to set a 2-D fine grid in both frequency domain and spatial domain, where two continuous parameters are needed to be discretized. Therefore, the computational burden of the former is much less than that of the latter.

Then, a least squares (LS) solution $\overline{\mathbf{S}}^S(f)$ is given by

$$\overline{\mathbf{S}}^S(f) = \mathbf{G}_S^{\dagger}\mathbf{Y}(f), f \in \mathcal{F}. \qquad (27)$$

Because $\mathbf{A}$ and $\mathbf{B}_{\Omega}$ are full rank matrices according to the discussions on the uniqueness of trilinear decomposition in Section 4.1, $\mathbf{G}_S = (\mathbf{A} * \mathbf{B}_{\Omega})$ is also a full rank matrix to get $\mathbf{G}_S^{\dagger} = \left(\mathbf{G}_S^{\mathrm{H}}\mathbf{G}_S\right)^{-1}\mathbf{G}_S^{\mathrm{H}}$.

Similarly, f_k and θ_k can be calculated through the steps (d)-(f) of Table 1. Therefore, the second method named joint algorithm based on subspace decomposition (JDFSD) is given in Table 2.

Table 2
Algorithm JDFSD.

Step	Operation
(a)	Calculate **R** according to (24);
(b)	Gain $\mathbf{U}_N$ by applying SVD to **R**;
(c)	Compute $P(l, \phi)$ through (26);
(d)	Acquire ϕ_k, l_k by peak search algorithm, further, we have **A**, Ω, and $\mathbf{G}_S$;
(e)	Determine $\overline{\mathbf{S}}^S(f)$ according to (27);
(f)	Execute the step (d)–(f) of Table 1.

5. Performance analysis

5.1. Estimation performance

Basing on our proposed methods in Section 4, we can obtain the estimates of the spatial phase, carrier frequency, and DOA as well. From (2), we observe that a spatial phase contains the frequency and the DOA. The performance of spatial phase estimation is also discussed. It is worth noting that Ω is a set of discrete parameters because it is the set of frequency band positions as well as the support set of **B**. It is well known that if the wrong set of Ω is used, we cannot obtain the estimates of true values of spatial phase, frequency and DOA. It makes nonsense to discuss the performance and CRB of those estimates in this case. Therefore, it is reasonable to derive the CRB of the estimated parameter provided that Ω is known or can be previously estimated by the methods shown in Section 4, which is similar to the case of the derivation for the CRB of DOA in the array signal processing literatures, such as [51]. In a nutshell, the following derivations are based on the condition with known Ω, and $\mathcal{S}$ is also known due to the fact that the support set $\mathcal{S}$ is determined by Ω in (14).

5.1.1. Spatial phase estimation performance

In this section, we derive the CRB of spatial phase ϕ estimation based on our proposed signal model.

If the signal autocorrelation matrices are defined as

$$\mathbf{R_S} \triangleq \frac{1}{T} \sum_{f=1}^{T} \mathbf{S}(f)\mathbf{S}^H(f), \tag{28}$$

$$\mathbf{R_{\overline{S}}} \triangleq \frac{1}{T/L} \sum_{f=1}^{T/L} \overline{\mathbf{S}}^S(f)\overline{\mathbf{S}}^{SH}(f), \tag{29}$$

where $\overline{\mathbf{S}}^{SH}(f)$ denotes Hermitian transpose of $\overline{\mathbf{S}}^S(f)$. Based on the property of $\overline{\mathbf{S}}^S(f)$, it is easy to verify that

$$\mathbf{R_S} = \frac{1}{L}\mathbf{R_{\overline{S}}}. \tag{30}$$

According to the model of (12), the log-likelihood function of the data $\mathbf{Y}(f)$, $f \in \mathcal{F}$ is given by

$$\mathcal{L} = \text{const} - \frac{1}{\sigma^2} \sum_{f=1}^{T/L} \left(\mathbf{Y}(f) - \mathbf{G}_S\overline{\mathbf{S}}^S(f)\right)^H \cdot \left(\mathbf{Y}(f) - \mathbf{G}_S\overline{\mathbf{S}}^S(f)\right). \tag{31}$$

Making use of the APPENDIXE (E.1) and the conclusion of Section IV Eq. (4.6) in [51], the CRB of our model is given by

$$\begin{aligned}
\text{CRB}_{\text{sub}}(\phi) &= \frac{\sigma^2}{2T/L}\left(\Re\left(\left(\mathbf{E}^H\mathbf{P}_{\mathbf{G}_S}\mathbf{E}\right) \odot \mathbf{R}_{\overline{\mathbf{S}}}^H\right)\right)^{-1} \\
&= \frac{\sigma^2}{2T}\left(\Re\left(\left(\mathbf{E}^H\mathbf{P}_{\mathbf{G}_S}\mathbf{E}\right) \odot \mathbf{R}_{\mathbf{S}}^H\right)\right)^{-1}
\end{aligned} \tag{32}$$

where $\mathbf{P}_{\mathbf{G}_S} = \mathbf{I} - \mathbf{G}_S\mathbf{G}_S^\dagger$, $\mathbf{E} = [\mathbf{E}_1, \ldots, \mathbf{E}_K]$, $\mathbf{E}_i = \frac{d\mathbf{G}_{S_i}}{d\phi_i}$.

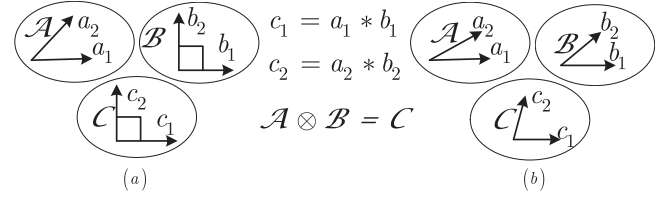The CRB based on the Nyquist sampling is denoted as $\text{CRB}_{\text{Ny}}(\phi))$ [51]

Fig. 2. 2-D unionized space.

$$\text{CRB}_{\text{Ny}}(\phi) = \frac{\sigma^2}{2T}\left(\Re\left(\left(\mathbf{D}^H\mathbf{P}_\mathbf{A}\mathbf{D}\right) \odot \mathbf{R}_\mathbf{S}^H\right)\right)^{-1} \tag{33}$$

where $\mathbf{P}_\mathbf{A} = \mathbf{I} - \mathbf{A}\mathbf{A}^\dagger$, $\mathbf{D} = [\mathbf{D}_1, \ldots, \mathbf{D}_K]$, $\mathbf{D}_i = \frac{d\mathbf{A}_i}{d\phi_i}$.

Next, we will show that the $\text{CRB}_{\text{sub}}(\phi)$ is lower than the $\text{CRB}_{\text{Ny}}(\phi)$ when the branch number of our architecture is equal to the factor L, i.e., $P = L$. In other words, the received data on the sub-Nyquist sampling can be though of as the rearrangement of the data on the Nyquist sampling.

$$\begin{aligned}
\text{CRB}_{\text{sub}}(\phi) &= \text{CRB}_{\text{Ny}}(\phi), \text{ when K} = 1, \\
\text{CRB}_{\text{sub}}(\phi) &\leq \text{CRB}_{\text{Ny}}(\phi), \text{ when } K > 1.
\end{aligned} \tag{34}$$

Proof: see Section 8. □

Remark 4. From the physical point of view, it is easy to obtain the cross-correlation δ_{ij} between $\mathbf{B}^i$ and $\mathbf{B}^j$, i.e., $\delta_{ij} = \left|\left(\mathbf{B}^i\right)^H\mathbf{B}^j\right|$ for $i \neq j$

$$\begin{cases} \delta_{ij} < 1, P < L \\ \delta_{ij} = 0, P = L \end{cases} \tag{35}$$

The degree of correlation of the new array steering vectors $\mathbf{G}_S$ is lower than that of the primary array steering vectors **A** provided that both (45) and (35) are guaranteed. For example, the new array steering vectors are uncorrelated when $P = L$ in spite of the primary array steering vectors are correlated. Making use of this property, the proposed method for estimating multiple DOAs can be considered as a problem for single DOA estimation. This is a reason why the performance of DOA estimation of our model is more robust than that of the primary model when the number of signals is increased.

From the geometrical point of view, the vectors $\mathcal{C}$ are orthogonal in space domain as long as the vectors $\mathcal{B}$ are orthogonal even if they are the same, as illustrated in Fig. 2 (a). Particularly, the degree of correlation of $\mathcal{C}$ is lower than that of both $\mathcal{A}$ and $\mathcal{B}$ for $P < L$, as depicted in Fig. 2 (b). Intuitively, the correlativity of array steering vectors accounts for the similarity and identifiability of the DOA, where the lower it is, the better the estimation performance is.

5.1.2. Frequency estimation performance

Each CRB value of multiple signals is the same as the CRB value for single-tone frequency estimation because enough snapshots approaches the correlativity to zero. Besides, the SNR can be improved for frequency estimation by coherent integration of multichannel data. Recall those two facts and the conclusion in [52], and the CRB_{Ny} and CRB_{sub} for frequency estimation are given as follow when T is sufficiently large.

$$\text{CRB}_{\text{Ny}}(f) = \frac{1}{4\pi^2} \frac{6}{\text{SNR}} \frac{1}{M} \frac{f_N^2}{T^3}. \tag{36}$$

$$\begin{aligned}
\text{CRB}_{\text{sub}}(f) &= \frac{1}{4\pi^2} \frac{6}{\text{SNR}} \frac{1}{MP} \frac{(f_N/L)^2}{(T/L)^3}. \\
&= \frac{1}{4\pi^2} \frac{6}{\text{SNR}} \frac{1}{M} \frac{L}{P} \frac{f_N^2}{T^3}.
\end{aligned} \tag{37}$$

It is easy to get that $\text{CRB}_{\text{sub}}(f) \geq \text{CRB}_{\text{Ny}}(f)$ since $L \geq P$.

5.1.3. DOA Estimation performance

From (2), we observe that the DOA depends on both the spatial phase and operating frequency. The phase estimation is much worse than the frequency estimation because the sampling number M in space domain is much less than the sampling number T in time domain. Herein, we assume that the DOA estimation precision is constrained by the spatial phase estimation ignoring the frequency factor so that the $CRB_{sub}(\theta)$ has the same expression as (32) except for the derivative of the matrix $\mathbf{G}_S$ with respect to θ. Recall the proof in Section 8, and a conclusion similar to (34) can be obtained.

5.2. Computational complexity

In the following analysis, we should assume $T \gg L$, $T \gg M$, $T \gg K$, $L \geq P \geq K$, $M > K$. The complexity of JDFTD algorithm is dominated by steps (a) and (d) in Table 1. The complexity of step (a) is ALS or RALS iterative operation, which costs respectively $(7(M + P) + 3MP)K\frac{T}{L}$ or $2K^2(M + P)\frac{T}{L}$ multiplications per iteration [53]. The main complexity of step (d) is IFFT operation and frequency MLE. The former costs $K \log_2 \frac{T}{L} \frac{T}{L}$ multiplications and the latter needs $\frac{T}{L}$ multiplications per frequency point each signal. The frequency MLE accuracy depends on fineness of searching grid. In order to reduce the computational complexity, we can explore the idea of adaptively refining the searching grid when implementing MLE according to the style in [54]. So frequency MLE needs several $K\frac{T}{L}$ multiplications. Therefore, the complexity of JDFTD algorithm is dominated by ALS or RALS iterative operation, which costs respectively $(7(M + P) + 3MP)K\frac{T}{L}$ or $2K^2(M + P)\frac{T}{L}$ multiplications per iteration.

The complexity of JDFSD algorithm is dominated by steps (b) and (c) in Table 2. The complexity of step (b) is SVD operation, which costs $\mathcal{O}(M^3P^3)$ multiplications [55]. The complexity of step (c) is the pseudo-spectra calculation, which costs M^2P^2L multiplications per spatial phase point. Searching for pseudo-spectra peaks can be divided into two steps: (i) Compute and find the K rough spatial phase in the range $(-\pi, \pi)$ by coarse grid. (ii) Compute and find more accurate spatial phase values near the previous values using adaptively refining the searching grid shown in [54]. Usually, the number of spatial phase points is from several hundreds to thousands. So, the complexity of JDFSD algorithm is dominated by pseudo-spectra calculation, which costs M^2P^2L multiplications per spatial phase point.

6. Simulation

In this section, we will evaluate the performance of the proposed algorithms using several numerical simulations. In these examples, we consider some non-coherent narrowband far-field signals with equal power impinging on a ULA composed of $M = 8$ sensors. The inter-sensor distance is a half wavelength in terms of the Nyquist sampling rate. The number of snapshots is fixed at $T = 4000$ for the Nyquist sampling and $T_{sub} = T/L$ for the sub-Nyquist sampling, where the Nyquist sampling rate is $f_N = 10$ GHz and the sampling rate reduction factor is $L = 20$.

Herein, the RMSE of DOA is defined as $RMSE = \sqrt{\frac{1}{N_m K} \sum_{i=1}^{N_m} \sum_{k=1}^{K} (\theta_k^i - \widehat{\theta}_k^i)^2}$, where the superscript i refers to the ith trial and N_m denotes the number of Monte Carlo tests. The RMSEs of spatial phase and frequency are defined similarly.

The receiver configuration parameters of ST-Euler-ESPRIT, i.e., sensor number M, branch number P, sampling rate reduction factor L, and minimum delay interval are the same with ours. In the following sections, we will illustrate the performances of the proposed algorithms and the ST-Euler-ESPRIT method with different noise levels, branch number and source number, respectively.

Table 3
Abbreviations

Abbreviation	Array	Algorithm/Low bound
CRB_{Ny}	ULA	CRB using Nyquist sampling
CRB_{sub}	ULA	CRB using sub-Nyquist sampling
JDFTD	ULA	algorithm listed in Table 1
JDFSD	ULA	algorithm listed in Table 2
ST-Euler-ESPRIT	ULA	algorithm in [36]
CRB_{Ny}(MRA)	MRA	CRB using Nyquist sampling
CRB_{sub}(MRA)	MRA	CRB using sub-Nyquist sampling
JDFTD(MRA)	MRA	algorithm listed in Table 1
JDFSD(MRA)	MRA	algorithm listed in Table 2

In the experiments, 20,000 Monte Carlo trials for each example are implemented. The simulations are performed using MATLAB 2016 B running on an Intel Core i7-2600 K, 3.40 GHz processor with 8 GB of memory, under Windows 10 64bit. The specifications of abbreviations are shown in Table 3.

Remark 5. For comparison, we briefly summarize the characteristic of the current schemes of [33–37]. (i) The number of sensors in [37] is more than that of [33–36] as well as our proposed scheme due to a L-shaped array used in [37]. (ii) In [33], the accurate estimation of the frequency and DOA is constrained by both the inverse of block length and array aperture. Besides, its computational burden is very large because a frequency-searching method is needed. (iii) If each channel is sampled at the same sub-Nyquist rate, the number of channels reflects the hardware complexity for [34–36]. The total number of channels is MP in [36], $2M$ in [34], $M + P$ in [35], and MP in our scheme, respectively. The receiver architecture in [36] is quite similar to the proposed one and the performance of ST-Euler-ESPRIT in [36] is better than that of in [34,35].

Given receiver architecture and implementation issues, we compare our methods with the ST-Euler-ESPRIT method [36].

6.1. Performance with various noise levels

In this section, we will study the performance of the proposed algorithms and the ST-Euler-ESPRIT method with respect to various noise levels. A MRA is considered to verify the array configuration flexibility of our receiver architecture and algorithms, where the MRA is composed of $M = 8$ sensors, which are located at $\mathbf{d} = [0, 1, 4, 10, 16, 22, 28, 30]d$. However, the ST-Euler-ESPRIT method is constrained by a ULA. The branch number is set as $P = L$ and the average sampling rate $f_A = \frac{P f_N}{L}$ is equal to the Nyquist sampling rate. Consider $K = 3$ signals transmitted from $\boldsymbol{\theta} = [\theta_1, \theta_2, \theta_3]$, where θ_i, $i = 1, 2, 3$, satisfies an uniform distribution in the intervals $[-2.5 + 10(i - 2), 2.5 + 10(i - 2)]°$. The signal carrier frequencies $\mathbf{f} = [f_1, f_2, f_3]$ are uniformly distributed in [0.5,9.5]GHz. Based on the previous assumption, signals should be distributed in the different sub-band.

Figs. 3–5 illustrate the RMSE curves versus SNR in terms of spatial phase, frequency, and DOA estimation, respectively. From Fig. 3, we can see that the phase estimation performance of the JDFTD and JDFSD improves when the SNR is increased and the curves are agreed with those of the CRB_{sub} for both the ULA and MRA. Since the ST-Euler-ESPRIT method is invalid when the MRA is employed, only the performance that employs ULA is given in this simulation. The estimation performance of the ST-Euler-ESPRIT method is inferior to that of the JDFTD and JDFSD methods. This is because that the data from delay channels is modeled in terms of the temporal steering vectors correlated when the number of channels is finite in [36]. Fortunately, by the compressed sampling technique, the column vectors of $\mathbf{B}$ of our model are constructed from uncorrelated frequencies for finite channels. We can also observe that
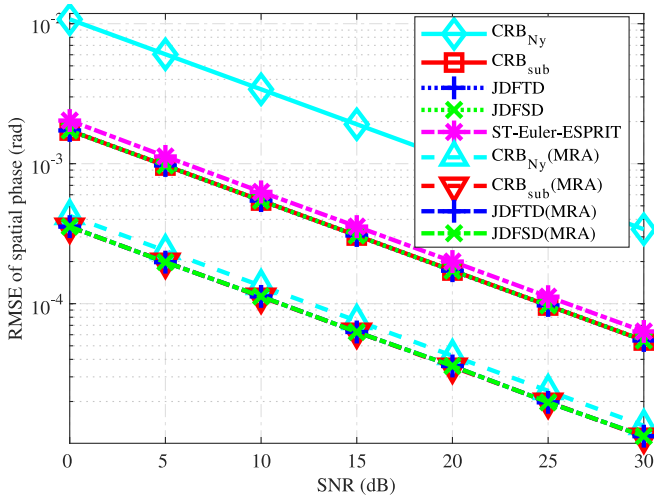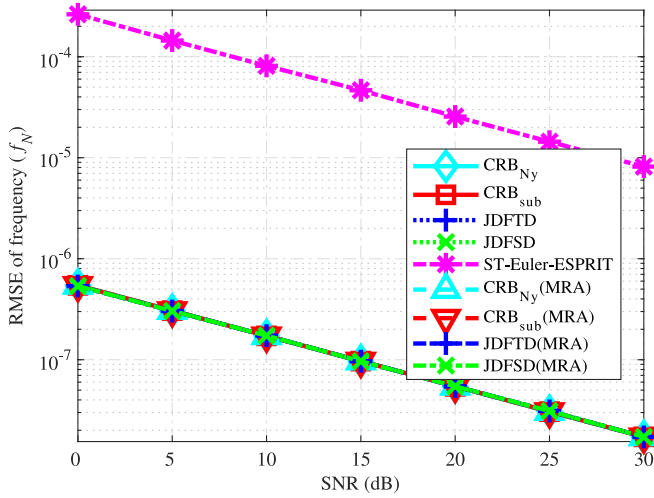
Fig. 3. RMSE of phase estimates versus SNR.

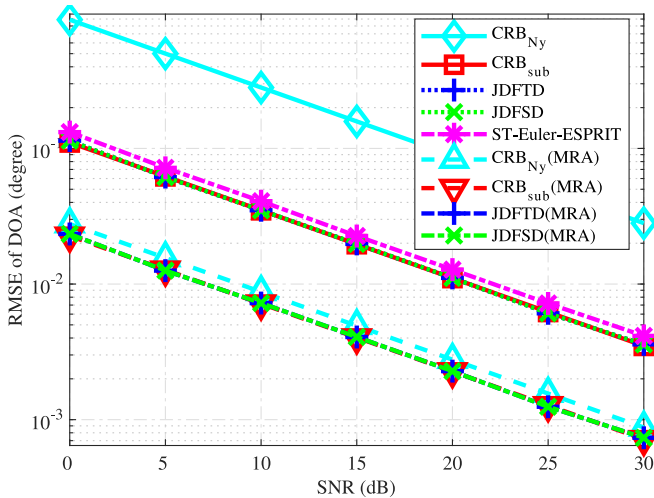**Fig. 4.** RMSE of frequency estimates versus SNR.

Fig. 5. RMSE of DOA estimates versus SNR.

Table 4
Averaged CPU run time.

Algorithm	Time(s)
JDFTD	9.6E–2
JDFSD	1.1E+0
ST-Euler-ESPRIT	1.7E–2
JDFTD(MRA)	8.3E–2
JDFSD(MRA)	6.9E–1

the JDFTD and JDFSD methods can significantly improve the phase estimation performance if a MRA is used. This is because that the MRA can extend the array aperture. Furthermore, the CRB_{sub} is lower than the CRB_{Ny} when $L = P$ and $K > 1$ as shown in Fig. 3. This illustrates that the proposed model can obtain a better phase estimation than the Nyquist sampling structure. The CRB_{Ny} suffers worst performance because the CRB_{Ny} does not involve the information in frequency domain in comparison to other methods.

Fig. 4 demonstrates the frequency estimation performance with respect to the SNR. The performance of ST-Euler-ESPRIT is remarkably worse than that of the JDFTD and JDFSD methods, which approach the CRB_{sub} for both the ULA and MRA structure. The reason is that the ST-Euler-ESPRIT for frequency estimation only makes use of the temporal steering vectors without using the signal samplings.

Comparing Fig. 5 with Fig. 3, it can be found that the performance of DOA estimation and phase estimation have the similar feature. Because the sampling number M in space domain is much less than the sampling number T in time domain, the precision of the phase estimation is worse than that of the frequency estimation. Further, the performance of DOA estimation is mainly constrained by the phase estimation. Thus, only the phase estimation results rather than the DOA estimation results is provided in the following simulations.

Table 4 presents the CPU run time of the proposed algorithms and the ST-Euler-ESPRIT method. Each of the time is the average result with different noise levels. From Table 4, we can see that the ST-Euler-ESPRIT costs the least due to the fact that the computational complexity of the ST-Euler-ESPRIT is dominated by the implementation of the SVD without searching or iterative procedure.

6.2. Performance with various branch number

In this section, we will investigate the estimation performance in the case of different branch number. The simulation parameters are the same with that in Section 6.1 except that the branch number P changes from 6 to 20 with an equal interval 2 - that is, the average sampling rate $f_A = \frac{P f_N}{L}$ changes from $0.3 f_N$ to f_N at $0.1 f_N$ interval.

In [26], **B** is a fully Kruskal-rank when L is a prime number. However, L is a compound number in our simulation. The sampling pattern is selected by checking whether **B** is a fully Kruskal-rank matrix according to the definition of Kruskal-rank. The practical procedure is as follows: (i) Generate a random sampling pattern C, i.e., select P numbers from the set $[0, 1, \ldots, L − 1]$ as candidate sampling pattern. (ii) Calculate corresponding matrix **B**. (iii) Select P columns from **B** to constitute $\mathbf{B}_c$, and judge whether $\mathbf{B}_c$ is a full rank matrix (an alternative method is to investigate the singularity of $\mathbf{B}_c$ through computing the condition number). If $\mathbf{B}_c$ is not a full rank matrix, the candidate sampling pattern C does not satisfy and return to the step (i). Otherwise, continue to check all possible ways. The candidate sampling pattern C is acceptable if $\mathbf{B}_c$ is a full rank matrix for all cases. We observe that the number of all possible ways is a combinatorial problem and significantly increases with L. However, there are two aspects worth noting that **B** is a

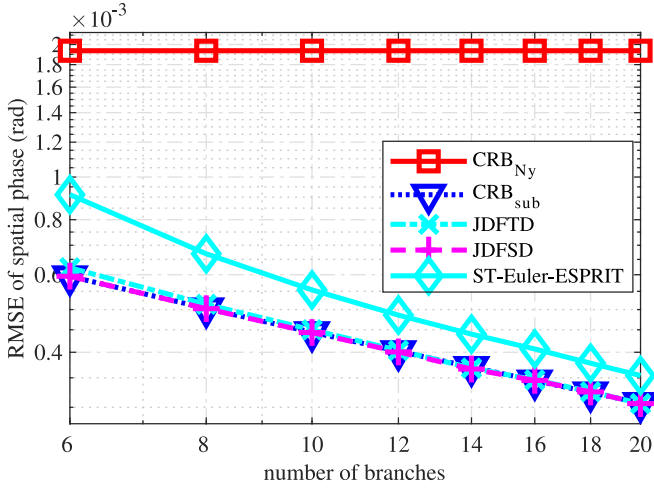**Fig. 6.** RMSE of phase estimates versus number of branches.

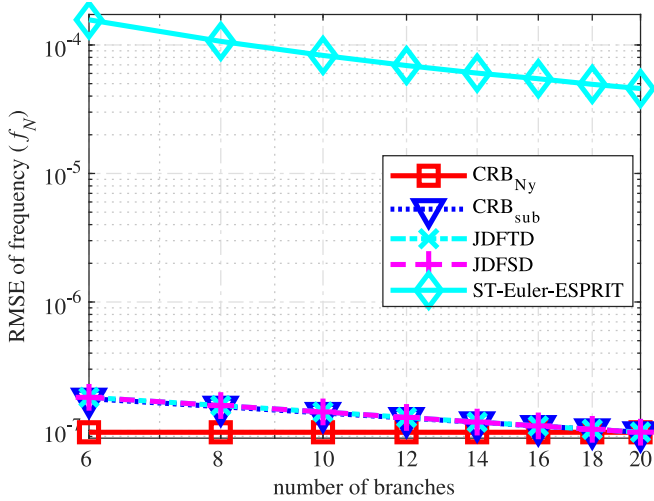**Fig. 8.** RMSE of phase estimates versus number of source signals.

Fig. 7. RMSE of frequency estimates versus number of branches.

fully Kruskal-rank matrix provided that L is a prime number and it is not necessary to search out all sampling patterns because only one satisfied sampling pattern is enough to design receiver structure in practical applications.

Fig. 6 shows that the phase (DOA) estimation performances of the JDFTD and JDFSD improve with the increasing of the branch number and reach the CRB$_{sub}$ when P is large enough. Besides, it can be seen that the performance of the JDFTD and JDFSD methods is superior to that of the ST-Euler-ESPRIT with respect to branch numbers, especially for the small ones. Meanwhile, it is observed that the performance of the JDFTD is slightly worse than that of the JDFSD and CRB$_{sub}$. This is because that the degree of correlation of the column vectors of **B** becomes very high for small P, which results in slow convergence of the trilinear decomposition. It is worth noting that the CRB$_{sub}$ is lower than the CRB$_{Ny}$ not only when $L = P$ but also when $P = 0.3 L$. This illustrates that decreasing the correlativity of array steering vectors from (35) can obtain more performance improvement compared to decreasing of samplings.

From Fig. 7, we can see that the performance of the JDFTD and JDFSD methods improves when the branch number is reduced and approaches to the CRB$_{sub}$ for large enough P. We can also observe from Figs. 6 and 7 that the CRB of the proposed model is lower than the traditional CRB in terms of DOA or spatial phase esti-
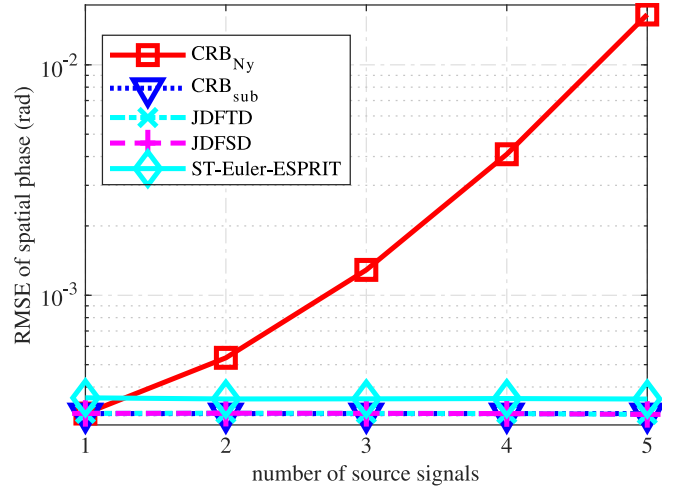
mation when the average sampling rate is lower than the Nyquist sampling rate, whereas the contrary results happen for frequency estimation. Generally, the estimation performance of DOA or spatial phase is much more worse than that of frequency because of $M \ll T$. Therefore, the improvement of estimation performance of DOA or spatial phase is more significant, which means that the degradation of estimation performance of frequency is negligible.

6.3. Performance with various signal number

In this section, we demonstrate the estimation performance with respect to the number of signals. The simulation parameters are the same as those of 6.1 except for the number of signals K changing from 1 to 5 with an equal interval 1. Assume that $\boldsymbol{\vartheta} = [\vartheta_1, \vartheta_2, \ldots, \vartheta_5]$, where ϑ_i, $i = 1, 2, \ldots, 5$, is randomly chosen from $[10j - 2.5, 10j + 2.5]°$, where $j = -2, -1, \ldots, 2$ and j is independent with i. Let $\boldsymbol{\upsilon} = [\upsilon_1, \upsilon_2, \ldots, \upsilon_5]$ be subject to [0.5,9.5] GHz uniformly distribution, and any two signals be not in the same sub-band at a fixed time. The DOAs $\boldsymbol{\theta} = [\vartheta_1, \vartheta_2, \ldots, \vartheta_K]$ and signal carrier frequencies $\mathbf{f} = [\upsilon_1, \upsilon_2, \ldots, \upsilon_K]$ vary with the signal number K.

From Fig. 8, we can see that the phase (DOA) estimation performance of the JDFTD, JDFSD and ST-Euler-ESPRIT is robust to the number of source signals. The RMSE curves of the JDFTD and JDFSD are agreed with that of the CRB$_{sub}$, whereas that of ST-Euler-ESPRIT is a little higher. Fig. 8 also shows that the CRB$_{Ny}$ is equal to the CRB$_{sub}$ only when $K = 1$ and increases with the signal number, which matches the analysed results in Section 5.1, where the increasing is even faster than the exponential function intuitively. Herein, the robustness to the number of source signals benefits from the orthogonality among the base vectors corresponding to different frequencies of our methods and the CRB$_{sub}$. As explanation in Section 6.1, the performance of ST-Euler-ESPRIT is worse than those of our methods due to the use of the space-time steering vectors.

Fig. 9 shows that the frequency estimation performance of the JDFTD , JDFSD and ST-Euler-ESPRIT is robust to the number of source signals as well. However, the ST-Euler-ESPRIT performs remarkably worse than the JDFTD and JDFSD, whose RMSEs can reach the CRB$_{sub}$. The reason can be seen in Section 6.1. Differing from the result in Fig. 8, the CRB$_{Ny}$ is equal to the CRB$_{sub}$ in Fig. 9 for $P = L$.
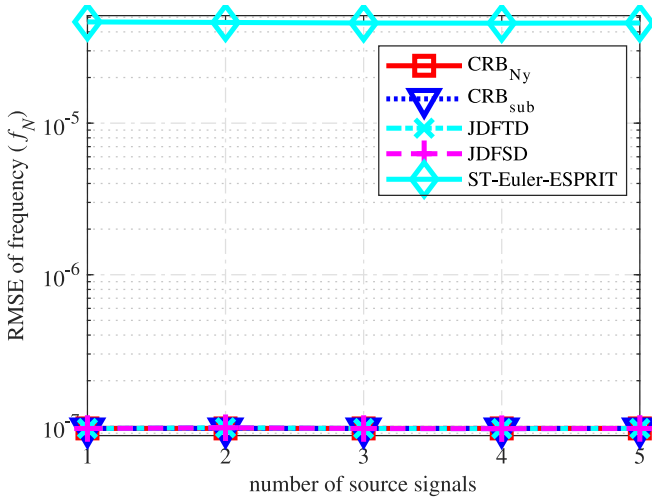
Fig. 9. RMSE of frequency estimates versus number of source signals.

7. Conclusions

In this paper, an array receiver architecture is designed for multiple narrowband far-field signals whose carrier frequencies are far separated, where the sub-Nyquist sampling technique is used. We derived a time-space unionized signal reception model by taking the spatial sampling and sub-Nyquist sampling into consideration simultaneously. Meanwhile, we can decrease the time-domain sampling rate and improve the DOA estimation performance.

Two joint DOA and frequency estimation algorithms are presented for such a model, one is based on trilinear decomposition in terms of third-order tensor and the other is based on subspace decomposition. The estimation performance and computational complexity of the proposed algorithms are analyzed and the CRBs of the spatial phase estimation, frequency estimation, and DOA estimation are derived for the model as well. In terms of spatial phase estimation, we proved that the CRB is immune to the signal number when the branch number of our architecture is equal to the sampling rate reduction factor, and is lower than the CRB for the conventional model which employs Nyquist sampling. Furthermore, the new array steering vectors are completely uncorrelated under the limited number of sensors, which makes a big improvement for the spatial phase estimation performance. From the geometry perspective, the estimation performance improvement benefits from the unionized time-space model.

The simulations validated that the receiver architecture and the proposed approaches are feasible, and the RMSEs of the proposed approaches are very close to their CRB and are beyond the CRB which employs Nyquist sampling in the case of different noise levels, different branch number, or different source number. Specifically, the variances of the proposed approaches are lower than the CRB which employs Nyquist sampling when the branch number of our architecture is less than the sampling rate reduction factor-that is, the average sampling rate is lower than the Nyquist sampling rate.

Appendix

Now we prove that the $\mathrm{CRB}_{\mathrm{sub}}(\phi)$ is lower than the $\mathrm{CRB}_{\mathrm{Ny}}(\phi)$ when the branch number of our architecture is equal to the sampling rate reduction factor ($P = L$). The sub-Nyquist sampling and Nyquist sampling obtain equal snapshots in the same time and the received data by sub-Nyquist sampling can just be viewed as the rearrangement of the received data by Nyquist sampling. The proof will carry out in two steps: first, the $\mathrm{CRB}_{\mathrm{Ny}}(\phi)$ increases with the

number of sources K, and then the $\mathrm{CRB}_{\mathrm{sub}}(\phi)$ is not influenced by the number of sources, and then the $\mathrm{CRB}_{\mathrm{sub}}(\phi)$ is equal to the minimum value of the $\mathrm{CRB}_{\mathrm{Ny}}(\phi)$.

Next, let's present a proposition for further derivations.

Proposition 1. *For a Hermitian matrix* $\mathbf{M}_+ \in \mathbb{C}^{(K+1)(K+1)}$, *if* $\mathbf{M}_+ \succeq \mathbf{0}$, *then* $\left(\mathbf{M}_+^{-1}\right)_K \succ ((\mathbf{M}_+)_K)^{-1}$.

Proof. Since $\mathbf{M}_+ \succeq 0$ it can be decomposed as $\mathbf{M}_+ = \mathbf{H}_+\mathbf{H}_+^{\mathrm{H}}$, where $\mathbf{H}_+ \in \mathbb{C}^{K \times K}$ [56]. Dividing $\mathbf{H}_+$ into blocks as $\mathbf{H}_+ = [\mathbf{H}, \mathbf{h}]$ results in

$$\mathbf{M}_+ = \begin{bmatrix} \mathbf{H}^{\mathrm{H}}\mathbf{H} & \mathbf{H}^{\mathrm{H}}\mathbf{h} \\ \mathbf{h}^{\mathrm{H}}\mathbf{H} & \mathbf{h}^{\mathrm{H}}\mathbf{h} \end{bmatrix} = \begin{bmatrix} \mathbf{M} & \mathbf{H}^{\mathrm{H}}\mathbf{h} \\ \mathbf{H}^{\mathrm{H}}\mathbf{h} & \mathbf{h}^{\mathrm{H}}\mathbf{h} \end{bmatrix}. \tag{38}$$

Recall the matrix inversion lemma [56], we can get that

$$\left(\mathbf{M}_+^{-1}\right)_K = \mathbf{M}^{-1} + \mathbf{V} \succeq \mathbf{M}^{-1} = ((\mathbf{M}_+)_K)^{-1} \tag{39}$$

because of $\mathbf{V} = \frac{1}{\mathbf{h}^{\mathrm{H}}(\mathbf{I}-\mathbf{H}\mathbf{H}^{\dagger})\mathbf{h}} \left(\mathbf{M}^{-1}\mathbf{H}^{\mathrm{H}}\mathbf{h}\right)^{\mathrm{H}} \left(\mathbf{M}^{-1}\mathbf{H}^{\mathrm{H}}\mathbf{h}\right) \succeq 0$. $\square$

For simplicity, let us introduce the following notations: $\mathbf{A}_+ = [\mathbf{A}, \mathbf{a}]$, $\mathbf{D}_+ = [\mathbf{D}, \mathbf{d}]$, $\mathbf{T}_+ = \mathbf{D}_+^{\mathrm{H}}\left(\mathbf{I} - \mathbf{A}_+\mathbf{A}_+^{\dagger}\right)\mathbf{D}_+$, $\mathbf{T} = \mathbf{D}^{\mathrm{H}}\left(\mathbf{I} - \mathbf{A}\mathbf{A}^{\dagger}\right)\mathbf{D}$, where $\mathbf{a}$ is the array steering vector corresponding to the increased angle ϕ_{K+1}, $\mathbf{d} = \frac{d\mathbf{a}}{d\phi_{K+1}}$. $\mathbf{R}_+ = \begin{bmatrix} \mathbf{R}_\mathbf{S} & \boldsymbol{\mu} \\ \boldsymbol{\mu} & \nu \end{bmatrix}$ is the covariance matrix of all $K + 1$ signals, where $\mathbf{R}_\mathbf{S}$ is the covariance matrix of original K signals, $\boldsymbol{\mu}$ is cross-correlation vector between the increased signal and the original signals, ν is average power of the increased signal. $(\mathbf{R})_K$ denotes the Kth order principal minor determinant of $\mathbf{R}$.

Making use of the nested structure of $\mathbf{A}_+$ and the matrix inversion lemma [56], we can obtain

$$\mathbf{A}_+\mathbf{A}_+^{\dagger} = \mathbf{A}\mathbf{A}^{\dagger} + \mathbf{U}, \tag{40}$$

where $\mathbf{U} = \frac{1}{\mathbf{a}^{\mathrm{H}}(\mathbf{I}-\mathbf{A}\mathbf{A}^{\dagger})\mathbf{a}}((\mathbf{I} - \mathbf{A}\mathbf{A}^{\dagger})\mathbf{a})((\mathbf{I} - \mathbf{A}\mathbf{A}^{\dagger})\mathbf{a})^{\mathrm{H}} \succeq \mathbf{0}$ because $(\mathbf{I} - \mathbf{A}\mathbf{A}^{\dagger})$ is a projection matrix. Taking (40) and the nested structure of $\mathbf{D}_+$ into $\mathbf{T}_+$ leads to

$$(\mathbf{T}_+)_K = \mathbf{T} - \mathbf{D}^{\mathrm{H}}\mathbf{U}\mathbf{D}. \tag{41}$$

It is easy to prove that $\mathbf{D}^{\mathrm{H}}\mathbf{U}\mathbf{D} \succeq 0$ with $\mathbf{U} \succeq 0$. We thus get

$$(\mathbf{T}_+)_K = \mathbf{T} - \mathbf{D}^{\mathrm{H}}\mathbf{U}\mathbf{D} \preceq \mathbf{T}. \tag{42}$$

Note that $\mathbf{T}_+ = \mathbf{D}_+^{\mathrm{H}}(\mathbf{I} - \mathbf{A}_+\mathbf{A}_+^{\dagger})\mathbf{D}_+ \succeq \mathbf{0}$, $\mathbf{R}_+ \succeq \mathbf{0}$, $\Re(\mathbf{T}_+ \odot \mathbf{R}_+^{\mathrm{H}}) \succeq \mathbf{0}$ holds. Making use of Proposition 1, we then obtain

$$\begin{aligned} \left((\Re(\mathbf{T}_+ \odot \mathbf{R}_+))^{-1}\right)_K &\succeq (\Re(\mathbf{T}_+ \odot \mathbf{R}_+)_K)^{-1} \\ &= (\Re((\mathbf{T}_+)_K \odot \mathbf{R}_\mathbf{S}))^{-1} \\ &\succeq (\Re(\mathbf{T} \odot \mathbf{R}_\mathbf{S}))^{-1}. \end{aligned} \tag{43}$$

From $\mathrm{CRB}_{\mathrm{Ny}+}(\phi) = \frac{\sigma^2}{2T}\left(\Re\left(\mathbf{T}_+ \odot \mathbf{R}_+^{\mathrm{H}}\right)\right)^{-1}$, $\mathrm{CRB}_{\mathrm{Ny}}(\phi) = \frac{\sigma^2}{2T}\left(\Re\left(\mathbf{T} \odot \mathbf{R}_\mathbf{S}^{\mathrm{H}}\right)\right)^{-1}$, we can get

$$\begin{aligned} \left(\mathrm{CRB}_{\mathrm{Ny}+}(\phi)\right)_K &= \frac{\sigma^2}{2T}\left(\left(\Re\left(\mathbf{T}_+ \odot \mathbf{R}_+^{\mathrm{H}}\right)\right)^{-1}\right)_K \\ &\succeq \frac{\sigma^2}{2T}\left(\Re\left(\mathbf{T} \odot \mathbf{R}_\mathbf{S}^{\mathrm{H}}\right)\right)^{-1} = \mathrm{CRB}_{\mathrm{Ny}}(\phi). \end{aligned} \tag{44}$$

Expression (44) shows that the estimate performance for $\phi = [\phi_1, \ldots, \phi_K]$ in the scene where there are only K signals from ϕ is better than that in the scene where there are both the K signals from ϕ and the increased signal from ϕ_{K+1}. In other words, the increase of signal number will degrade the estimate performance. It is also well known that the estimation variance is lowest when there is only one signal. After calculation, the lowest estimation

variance is $\frac{6}{\text{SNR}\cdot T\cdot M(M^2-1)}$, where signal-to-noise ratio (SNR) is defined as $\text{SNR} = (E(|\mathbf{s}(t)|^2)/\sigma^2)$.

However, the estimation performance based on the proposed model is not degraded to the increasing signal number. It's easy to get $\mathbf{E} = \mathbf{D} * \mathbf{B}_\Omega$ since $\mathbf{G}_S = \mathbf{A} * \mathbf{B}_\Omega$. We further hold that

$$\mathbf{G}_S^H \mathbf{G}_S = \left(\mathbf{A}^H \mathbf{A}\right) \odot \left(\mathbf{B}_\Omega^H \mathbf{B}_\Omega\right), \tag{45}$$

$$\mathbf{E}^H \mathbf{E} = \left(\mathbf{D}^H \mathbf{D}\right) \odot \left(\mathbf{B}_\Omega^H \mathbf{B}_\Omega\right). \tag{46}$$

Thus, $\mathbf{G}_S^H \mathbf{G}_S = \left(\mathbf{A}^H \mathbf{A}\right) \odot \mathbf{I}$, $\mathbf{E}^H \mathbf{E} = \left(\mathbf{A}^H \mathbf{A}\right) \odot \mathbf{I}$ hold when $P = L$. Based on above results and after some matrix manipulations, we know that the estimation variance based on our model maintains at $\frac{6}{\text{SNR}\cdot T\cdot M(M^2-1)}$ for $P = L$. Consequently, we have

$$\begin{aligned} \text{CRB}_{\text{sub}}(\phi) &= \text{CRB}_{\text{Ny}}(\phi), \quad \text{when } K = 1, \\ \text{CRB}_{\text{sub}}(\phi) &\leq \text{CRB}_{\text{Ny}}(\phi), \quad \text{when } K > 1. \end{aligned} \tag{47}$$

References

[1] S. Haykin, Cognitive radio: brain-empowered wireless communications, IEEE J. Sel. Areas Commun. 23 (2) (2005) 201–220.

[2] S.V. Nagaraj, Entropy-based spectrum sensing in cognitive radio, Signal Process. 89 (2) (2009) 174–180.

[3] T. Yucek, H. Arslan, A survey of spectrum sensing algorithms for cognitive radio applications, IEEE Commun. Surv. Tutor. 11 (1) (2009) 116–130.

[4] M. Mishali, Y.C. Eldar, Wideband spectrum sensing at sub-Nyquist rates, IEEE Signal Process. Mag. 28 (4) (2011) 102–135.

[5] A. Mohammadi, M.R. Taban, J. Abouei, H. Torabi, Fuzzy likelihood ratio test for cooperative spectrum sensing in cognitive radio, Signal Process. 93 (5) (2013) 1118–1125.

[6] H. Sun, A. Nallanathan, C.X. Wang, Y. Chen, Wideband spectrum sensing for cognitive radio networks: a survey, IEEE Wirel. Commun. 20 (2) (2013) 74–81.

[7] D. Cohen, Y.C. Eldar, sub-Nyquist sampling for power spectrum sensing in cognitive radios: a unified approach, IEEE Trans. Signal Process. 62 (15) (2014) 3897–3910.

[8] M. Shafi, A.F. Molisch, P.J. Smith, T. Haustein, P. Zhu, P.D. Silva, F. Tufvesson, A. Benjebbour, G. Wunder, 5G: a tutorial overview of standards, trials, challenges, deployment, and practice, IEEE J. Sel. Areas Commun. 35 (6) (2017) 1201–1221.

[9] I.F. Akyildiz, S. Nie, S.-C. Lin, M. Chandrasekaran, 5G Roadmap: 10 key enabling technologies, Comput. Netw. 106 (2016) 17–48.

[10] X. Ge, R. Zi, H. Wang, J. Zhang, M. Jo, Multi-user massive MIMO communication systems based on irregular antenna arrays, IEEE Trans. Wirel. Commun. 15 (8) (2016) 5287–5301.

[11] H. Krim, M. Viberg, Two decades of array signal processing research: the parametric approach, IEEE Signal Process. Mag. 13 (4) (1996) 67–94.

[12] R. Schmidt, Multiple emitter location and signal parameter estimation, IEEE Trans. Antennas Propag. 34 (3) (1986) 276–280.

[13] R. Roy, A. Paulraj, T. Kailath, ESPRIT–A subspace rotation approach to estimation of parameters of cisoids in noise, IEEE Trans. Acoust. Speech Signal Process. 34 (5) (1986) 1340–1342.

[14] J.-F. Gu, W.-P. Zhu, M. Swamy, Direction of arrival tracking for signals with known waveforms based on block least squares techniques, J. Frankl. Inst. 354 (11) (2017) 4573–4594.

[15] F. Izedi, M. Karimi, M. Derakhtian, Joint DOA estimation and source number detection for arrays with arbitrary geometry, Signal Process. 140 (2017) 149–160.

[16] K. Wang, J.F. Gu, F. Ren, K. Wu, A multitarget active backscattering 2-D positioning system with superresolution time series post-processing technique, IEEE Trans. Microw. Theory Tech. 65 (5) (2017) 1751–1766.

[17] C. Fernandez-Prades, J. Arribas, P. Closas, Robust GNSS receivers by array signal processing: theory and implementation, Proc. IEEE 104 (6) (2016) 1207–1220.

[18] W. Gao, S. Emaminejad, H.Y.Y. Nyein, S. Challa, K. Chen, A. Peck, H.M. Fahad, H. Ota, H. Shiraki, D. Kiriya, Fully integrated wearable sensor arrays for multiplexed in situ perspiration analysis, Nature 529 (7587) (2016) 509.

[19] J.-F. Gu, W.-P. Zhu, M. Swamy, Fast and efficient DOA estimation method for signals with known waveforms using nonuniform linear arrays, Signal Process. 114 (2015) 265–276.

[20] A.N. Lemma, A.J. van der Veen, E.F. Deprettere, Joint angle-frequency estimation using multi-resolution ESPRIT, in: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), vol. 4, 1998, pp. 1957–1960 vol.4.

[21] M. Mohammed, DOA And frequency estimation using fast subspace algorithms, Signal Process. 77 (1) (1999) 49–62.

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

[23] E.J. Candes, M.B. Wakin, An introduction to compressive sampling, IEEE Signal Process. Mag. 25 (2) (2008) 21–30.

[24] J.H.G. Ender, On compressive sensing applied to radar, Signal Process. 90 (5) (2010) 1402–1414.

[25] Y.C. Eldar, T. Michaeli, Beyond bandlimited sampling, IEEE Signal Process. Mag. 26 (3) (2009) 48–68.

[26] M. Mishali, Y.C. Eldar, Blind multiband signal reconstruction: compressed sensing for analog signals, IEEE Trans. Signal Process. 57 (3) (2009) 993–1009.

[27] L. Yu, H. Sun, J.P. Barbot, G. Zheng, Bayesian compressive sensing for cluster structured sparse signals, Signal Process. 92 (1) (2012) 259–269.

[28] S.A. Razavi, M. Valkama, D. Cabric, Covariance-based OFDM spectrum sensing with sub-Nyquist samples, Signal Process. 109 (2015) 261–268.

[29] Y. Gu, N.A. Goodman, S. Hong, Y. Li, Robust adaptive beamforming based on interference covariance matrix sparse reconstruction, Signal Process. 96 (B) (2014) 375–381.

[30] S. Qin, Y.D. Zhang, M.G. Amin, B. Himed, DOA estimation exploiting a uniform linear array with multiple co-prime frequencies, Signal Process. 130 (2017) 37–46.

[31] F. Xi, S. Chen, Y.D. Zhang, Z. Liu, Gridless quadrature compressive sampling with interpolated array technique, Signal Process. 133 (2017) 1–12.

[32] A. Chaturvedi, H.H. Fan, Wideband delay and direction of arrival estimation using sub-Nyquist sampling, Signal Process. 135 (2017) 67–80.

[33] D.D. Ariananda, G. Leus, Compressive joint angular-frequency power spectrum estimation, in: Proceedings of the European Signal Processing Conference (EUSIPCO), 2013, pp. 1–5.

[34] A. Anil Kumar, S.G. Razul, C.-M.S. See, Spectrum blind reconstruction and direction of arrival estimation of multi-band signals at sub-Nyquist sampling rates, Multidimens. Syst. Sig. Process. (2016).

[35] A.A. Kumar, S.G. Razul, C.M.S. See, Spectrum blind reconstruction and direction of arrival estimation at sub-Nyquist sampling rates with uniform linear array, in: Proceedings of the IEEE International Conference on Digital Signal Processing (DSP), 2015, pp. 670–674.

[36] A.A. Kumar, S.G. Razul, M.G. Chandra, C.M. See, P. Balamuralidhar, Joint frequency and direction of arrival estimation with space-time array, in: Proceedings of the IEEE Sensor Array and Multichannel Signal Process. Workshop (SAM), 2016, pp. 1–5.

[37] S.S. Ioushua, O. Yair, D. Cohen, Y.C. Eldar, Cascade: compressed carrier and DOA estimation, IEEE Trans. Signal Process. 65 (10) (2017) 2645–2658.

[38] M. Mishali, Y.C. Eldar, From theory to practice: sub-Nyquist sampling of sparse wideband analog signals, IEEE J. Sel. Top. Signal Process. 4 (2) (2010) 375–391.

[39] M. Mishali, Y.C. Eldar, Sub-Nyquist sampling, IEEE Signal Process. Mag. 28 (6) (2011) 98–124.

[40] C. Proukakis, A. Manikas, Study of ambiguities of linear arrays, in: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), vol. 4, IEEE, 1994, pp. IV–549.

[41] K.-C. Tan, G.-L. Oh, M. Er, A study of the uniqueness of steering vectors in array processing, Signal Process. 34 (3) (1993) 245–256.

[42] Y.I. Abramovich, O. Gray, A.Y. Gorokhov, N.K. Spencer, Comparison of DOA estimation performance for various types of sparse antenna array geometries, in: Proceedings of the European Signal Processing Conference (EUSIPCO), IEEE, 1996, pp. 1–4.

[43] P. Pal, P.P. Vaidyanathan, Nested arrays: a novel approach to array processing with enhanced degrees of freedom, IEEE Trans. Signal Process. 58 (8) (2010) 4167–4181.

[44] C.L. Liu, P.P. Vaidyanathan, Remarks on the spatial smoothing step in coarray music, IEEE Signal Process. Lett. 22 (9) (2015) 1438–1442.

[45] J.-F. Gu, W.-P. Zhu, M. Swamy, Joint 2-D DOA estimation via sparse l-shaped array, IEEE Trans. Signal Process. 63 (5) (2015) 1171–1182.

[46] J.B. Kruskal, Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics, Linear Algebra Appl. 18 (2) (1977) 95–138.

[47] L.D. Lathauwer, A link between the canonical decomposition in multilinear algebra and simultaneous matrix diagonalization, SIAM J. Matrix Anal. A. 28 (3) (2006) 642–666.

[48] L. De Lathauwer, D. Nion, Decompositions of a higher-order tensor in block terms-part III: alternating least squares algorithms, SIAM J. Matrix Anal. A. 30 (3) (2008) 1067–1083.

[49] C. Navasca, L.D. Lathauwer, S. Kindermann, Swamp reducing technique for tensor decomposition, in: Proceedings of the European Signal Processing Conference (EUSIPCO), 2008, pp. 1–5.

[50] N. Li, S. Kindermann, C. Navasca, Some convergence results on the regularized alternating least-squares method for tensor decomposition, Linear Algebra Appl. 438 (2) (2013) 796–812.

[51] P. Stoica, A. Nehorai, MUSIC, Maximum likelihood, and cramér-Rao bound: further results and comparisons, IEEE Trans. Acoust. Speech Signal Process. 38 (12) (1990) 2140–2150.

[52] D.C. Rife, R.R. Boorstyn, Single tone parameter estimation from discrete-time observations, IEEE Trans. Inf. Theory 20 (5) (1974) 591–598.

[53] P. Comon, X. Luciani, A.L.F.D. Almeida, Tensor decompositions, alternating least squares and other tales, J. Chemom. 23 (7–8) (2010) 393–405.

[54] D. Malioutov, M. Cetin, A. Willsky, A sparse signal reconstruction perspective for source localization with sensor arrays, IEEE Trans. Signal Process. 53 (8) (2005) 3010–3022.

[55] G.H. Golub, C.F. van Van Loan, Matrix Computations, The Johns Hopkins University Press, 1996.

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