RESEARCH   **Open Access**

2-D DOA tracking using variational sparse Bayesian learning embedded with Kalman filter

Qinghua Huang[*], Jingbiao Huang, Kai Liu and Yong Fang

Abstract

In this paper, we consider the 2-D direction-of-arrival (DOA) tracking problem. The signals are captured by a uniform spherical array and therefore can be analyzed in the spherical harmonics domain. Exploiting the sparsity of source DOAs in the whole angular region, we propose a novel DOA tracking method to estimate the source locations and trace their trajectories by using the variational sparse Bayesian learning (VSBL) embedded with Kalman filter (KF). First, a transition probabilities (TP) model is used to build the state transition process, which assumes that each source moves to its adjacent grids with equal probability. Second, the states are estimated by KF in the variational E-step of the VSBL and the variances of the state noise and measurement noise are learned in the variational M-step of the VSBL. Finally, the proposed method is extended to deal with the off-grid tracking problem. Simulations show that the proposed method has higher accuracy than VSBL and KF methods.

Keywords: 2-D direction-of-arrival (DOA) tracking, Spherical array, Transition probabilities (TP) model, Variational sparse Bayesian learning (VSBL), Kalman filter (KF)

1 Introduction

Direction-of-arrival (DOA) estimation is an active research field of array signal processing and has been used in various applications, such as radar, channel modeling, tracking, and surveillance [1–3]. Among the estimation algorithms, multiple signal classification (MUSIC) and estimation of signal parameters via rotational invariance techniques (ESPRIT) are the most representative methods, which employ the signal and noise subspaces. Compared with the conventional beamforming algorithms, these methods enhance the estimation precision. However, the number of impinging signals must be the prior knowledge and the computational complexity of decomposing the covariance matrix increases when the number of array elements rises. Recently, sparse reconstruction methods have attracted substantial attention because the signals impinging on an array are intrinsically sparse in the spatial domain [4, 5]. In these methods, the whole angular domain is divided into some predefined grids and a measurement matrix is constructed by sampling these grids.

Based on the singular value decomposition (SVD), an approach named as l_1-SVD was proposed to reduce the computational complexity and enforce sparsity using l_1-norm [6]. Compared with the l_1-SVD method, sparse Bayesian learning (SBL) can model the sparse signals more flexibly and give more accurate recovery results [7–9]. When SBL was used to estimate DOA, the sparse prior for the interested signals is Gaussian [10] or Laplacian distribution [11]. The SBL method can achieve good estimation results for static DOA estimation.

In the case of tracking moving targets, most methods assume that each source angle is constant during a time interval. However, it may differ from one interval to another because of the moving sources [12]. There are some different approaches to track the DOAs of moving sources, such as classical subspace optimization approach, sparsity recovery theory, and adaptive filter method. The subspace optimization approaches only optimize the signal or noise subspace without using the eigenvalue decomposition (EVD) so that they can reduce the computational complexity and storage requirements. Yang presented a new approach to track the signal subspace using an unconstrained minimization method [13]. The subspace method, which

* Correspondence: qinghua@shu.edu.cn
Key Laboratory of Specialty Fiber Optics and Optical Access Networks, Shanghai University, Shanghai 200072, People's Republic of China

Huang *et al. EURASIP Journal on Advances in Signal Processing* (2018) 2018:23

Page 2 of 14

can be used to dynamic DOAs, was extended in L-shaped array [14] and two parallel linear arrays [15], respectively. The performance of the subspace method relies on the number of snapshots. This kind of method is inapplicable for the moving targets when the number of snapshots in each time interval is relatively small.

Vaswani et al. reviewed many algorithms on the analysis of dynamic sparse signal recovery. If the support change is highly correlated and the correlation model is known, we can get an accurate support estimation by using the previous estimation information [16–18]. In [19, 20], these methods combined the slow signal value change and slow support change to enhance the precision. Most of these methods are the deformation of basis pursuit denoising. A sequential Bayesian algorithm was introduced to estimate the moving DOAs in the time-varying circumstance [21]. It assumed that the sources move at a constant velocity and used this hypothesis to construct the signal-moving model. The locally competitive algorithm (LCA) was proposed to construct a dynamic system and track DOAs [22]. This method mainly introduced a thresholding function to enforce sparsity. A model was proposed to describe a time-varying array response in the frequency domain for each source [23]. The key of this paper is that it utilizes a hidden Markov model to describe the moving signal and uses the posterior inference to estimate the signal positions. It traps into the local optimum when the signals alias at high frequency.

The Kalman filter (KF) method is the most representative method for tracking sources in the adaptive filter theory, which means the value of the predicted state should equal to the value of the actual state by minimizing the Bayesian mean square error of the state vector. However, the variance of the noise and state can be seen as prior in the KF method. In [24], the time difference of arrival (TDOA) information was calculated by a pair of microphones and a distributed unscented KF method was used to track speakers in a nonlinear measurement model. The method combining KF with compression sensing was also used for dynamic DOA estimation [25, 26]. The state transition function was built under the assumption that the bearing change rate had been known [25]. This assumption is hard to be satisfied in real applications. A Bayesian compressive sensing Kalman filter (BCSKF) method [26] was proposed to track dynamic moving sources. This method used the constant DOA changes in the KF prediction, which meant that the source moved to the designated direction with a fixed step. The particle filter was utilized to track the trajectory of target. It adopted a series of particles to represent the posterior distribution of signal stochastic process. For example, an algorithm combining compressive sense and particle filter was

introduced in [27]. However, it only used the compressive sense to estimate the original position and utilized particle filter to track the trajectory.

In this paper, a spherical array is used to track the moving sources in 3-D space because the array is 3D symmetric structure and can capture high-order sound field information [28–30]. Besides, comparing with linear array, spherical array can capture the two-dimension information of signal rather than single-angle information. In addition, due to the special construction of spherical array, the receiving signal can be unfolded in special harmonic domain, which can separate the signal position coordinates and sensor position coordinates conveniently.

A new method is proposed based on the spherical array to track the 2-D dynamic DOAs. It combines the variational Bayesian inference and KF to improve the tracking performance. First, a transition probabilities (TP) model is built to describe the state transition process of a signal. In this model, the source moves to an uncertain situation with an equal probability rather than a determining expected DOA change in [26]. In addition, TP model is convenient to build with a tracking framework and adopts sparse methods to estimate signal position. Based on this model, an alternating iterative method is developed to track DOAs. In the first step, we use the KF method to estimate the signal values. In the second step, we use variational sparse Bayesian learning (VSBL) to learn the variances of the measurement noise and state noise, which are useful to update the signal state in the KF method. This is an interdependent process.

There are three differences between the proposed approach and the method appeared in paper [26]. On the one hand, the proposed method constructs a real-valued steering matrix in signal model rather than splitting the complex value into several real-valued ones, which can reduce a half of the computational quantity approximately. On the other hand, the KF estimation values is used instead of the Bayesian estimation values. The main contribution of paper [26] is putting the estimation parameters of KF to optimize the low bound of relate vector machine. However, the KF estimation will be embedded in VSBL to improve the precision further in this paper. Finally, we introduce the off-grid model to overcome the mismatch problems.

The rest of the paper is organized as follows. The real-valued array signal model is given in Section 2. The variational Bayesian inference is briefly reviewed, and the proposed tracking method is introduced in Sections 3 and 4. Numerical examples and simulation results are given in Section 5. Section 6 concludes the paper.

The notations description as:

$(a, b, A, B, \cdots)$, scalar variable	$\arg(\cdot)$, phase operator
$(\mathbf{A}, \mathbf{B}, \cdots)$, the matrix variable	$\lvert \cdot \rvert$, the absolute value
$(\mathbf{a}, \mathbf{b}, \cdots)$, column vector	$(\cdot)^+$, the Moore-Penrose pseudo inverse
$(\cdot)^T$, transpose operator	$(\cdot)^{'}$, the derivation operator
$(\cdot)^*$, complex conjugation operator	$\lVert \cdot \rVert$, l_2 norm
$(\cdot)^H$, conjugate transpose operator	$\lVert \cdot \rVert_F$, Frobenius norm
$\mathrm{diag}(\cdot)$, diagonal matrix	$\mathbf{A}(n, :)$, the nth row of matrix $\mathbf{A}$
$\mathrm{blkdiag}(\cdot)$, block diagonal matrix	$\mathbf{A}(:, n)$, the nth column of matrix $\mathbf{A}$
$\mathrm{Re}(\cdot)$, the real parts of a complex value	
$\exp(\cdot)$, the exponent signal	$\mathrm{Im}(\cdot)$, the imaginary parts of a complex value
k, the wavenumber	D, the number of signals
(ϑ_l, φ_l), the elevation and azimuth of the sensor	$(\breve{\theta}_{d,t}, \breve{\phi}_{d,t})$, the elevation and azimuth of the dth signal at the tth time interval
L, the number of sensors	t, time interval
B, snapshots	R, the radius of sphere array
G_1, G_2, the azimuth and elevation range	i, the imaginary unit $\sqrt{-1}$
$(\breve{\mathbf{X}}, \overline{\mathbf{X}}, \mathbf{X})$, the space domain, spherical domain and real-valued receiving signal	$(\breve{\mathbf{A}}, \hat{\mathbf{A}}, \overline{\mathbf{A}}, \mathbf{A})$, the true steering, dictionary, the spherical and real-valued matrix
$(\breve{\mathbf{S}}, \overline{\mathbf{S}}, \mathbf{S})$, the space amplify, the sparse amplify signal and the real-valued amplify signal	$(\breve{\mathbf{V}}, \overline{\mathbf{V}}, \mathbf{V})$, the space domain, the spherical and the real-valued noise
h_n, spherical Hankel function of order n	$Y_n^m(\cdot)$, the spherical harmonic of order n and degree m
$\mathbf{I}_n$, an $n \times n$ identity matrix	j_n, the n-order spherical Bessel function
$\mathbf{J}_n$, the exchange matrix with ones on its antidiagonal and zeros elsewhere	$(\cdot)^{'}$, the derivation operators
$\Gamma(\cdot)$, a Gamma function	$\mathbf{0}_n$, a column vector containing n zeros
$(\mathbf{A} \backslash \mathbf{b})$, take the variables $\mathbf{A}$ except variable $\mathbf{b}$	

2 Real-valued array signal model

Assuming D dynamic narrowband far-field signals with the wavenumber k impinge on a spherical array from $\breve{\Phi}_{d,t} = (\breve{\theta}_{d,t}, \breve{\phi}_{d,t})$. It is assumed that there are B snapshots available to process the received data and estimate DOA in each time interval. We use t to represent the range $[(t-1)B+1, \cdots, tB]$ for simplicity. The signal angular change is slow so that it can be considered as fixed in the time interval t. The spherical array is made of L identical isotropic elements with radius R in Fig. 1. The position of the lth sensor can be described as $\mathbf{R}_l = R[\cos\varphi_l \sin\vartheta_l, \sin\varphi_l \sin\vartheta_l, \cos\vartheta_l]^T$. The steering vector of the array is defined as
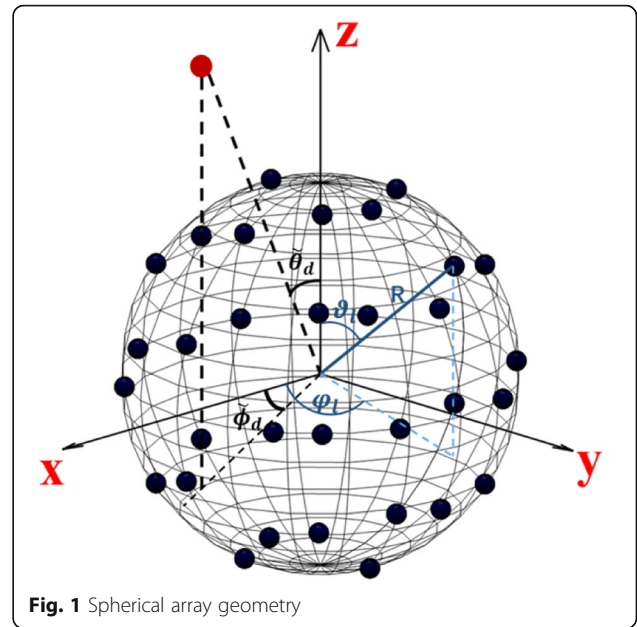

Fig. 1 Spherical array geometry

$$\breve{\mathbf{a}}\left(\breve{\theta}_{d,t}, \breve{\phi}_{d,t}\right) = \begin{bmatrix} \exp\left(ikR\left(\sin\breve{\theta}_{d,t}\sin\vartheta_1\cos\left(\breve{\phi}_{d,t}-\varphi_1\right)+\cos\breve{\phi}_{d,t}\cos\varphi_1\right)\right) \\ \vdots \\ \exp\left(ikR\left(\sin\breve{\theta}_{d,t}\sin\vartheta_L\cos\left(\breve{\phi}_{d,t}-\varphi_L\right)+\cos\breve{\phi}_{d,t}\cos\varphi_L\right)\right) \end{bmatrix}. \tag{1}$$

where $i = \sqrt{-1}$. The output of the array is given by

$$\breve{\mathbf{X}}_t = \breve{\mathbf{A}}_t \breve{\mathbf{S}}_t + \breve{\mathbf{V}}_t, \tag{2}$$

where $\mathbf{x}_t = [\breve{\mathbf{x}}_{1,t}, \cdots, \breve{\mathbf{x}}_{L,t}]^T$, $\mathbf{x}_{l,t} = [x_l((t-1)B+1), \cdots, x_l(tB)]^T$ is a column vector, which denotes the lth sensor receiving signal, $\breve{\mathbf{S}}_t = [\breve{\mathbf{s}}_{1,t}, \cdots, \breve{\mathbf{s}}_{D,t}]^T$ is the amplitude of source signal, $\breve{\mathbf{s}}_{d,t} = [s_d((t-1)B+1), \cdots, s_d(tB)]$. $\breve{\mathbf{A}}_t = [\breve{\mathbf{a}}(\breve{\theta}_{1,t}, \breve{\phi}_{1,t}), \cdots, \breve{\mathbf{a}}(\breve{\theta}_{D,t}, \breve{\phi}_{D,t})]$ is the steering matrix, and $\breve{\mathbf{V}}_t = [\breve{\mathbf{v}}_{1,t}, \cdots, \breve{\mathbf{v}}_{L,t}]^T$ is the measurement noise, and $\breve{\mathbf{v}}_{l,t} = [v_l((t-1)B+1), \cdots, v_l(tB)]^T$.

In order to construct a sparse signal model, we divide the azimuth and elevation range into $G_1 >> D$ and $G_2 >> D$ grids, respectively. That is to say, the whole 2-D angular space is divided into $G = G_1 G_2$ grids noted as $\mathbf{\Phi} = \{\Phi_1, \cdots, \Phi_G\}$, where $\Phi_g = (\theta_{g_1}, \phi_{g_2})$, $g_1 = 1, \cdots, G_1$, $g_2 = 1, \cdots, G_2$. The output of the sensors at the time interval t can be reformulated as

$$\breve{\mathbf{X}}_t = \hat{\mathbf{A}} \overline{\mathbf{S}}_t + \breve{\mathbf{V}}_t, \tag{3}$$

where $\hat{\mathbf{A}} = [\breve{\mathbf{a}}(\Phi_1), \cdots, \breve{\mathbf{a}}(\Phi_G)]$ is the redundant dictionary matrix containing the interested angles $\breve{\Phi}_{d,t}$, $d = 1, \cdots, D$, and $\overline{\mathbf{S}}_t = [\overline{\mathbf{s}}_{1,t}, \cdots, \overline{\mathbf{s}}_{G,t}]^T$ is a sparse matrix with only D nonzero rows corresponding to the positions of the sources. $\breve{a}_l(\Phi_g)$ denotes the lth element

Huang *et al. EURASIP Journal on Advances in Signal Processing* (2018) 2018:23

Page 4 of 14

of the dictionary vector $\breve{\mathbf{a}}\left(\Phi_g\right)$, and it can be represented using spherical harmonics [31] as

$$\breve{a}_l\left(\Phi_g\right) = \sum_{n=0}^{N} \sum_{m=-n}^{n} b_n(kR)\left[Y_n^m\left(\Phi_g\right)\right]^* Y_n^m(\Omega_l), \qquad (4)$$

where $\Omega_l = (\vartheta_l, \varphi_l)$, N is the maximum spherical harmonics order, $b_n(kR)$ is the far-field mode strength, which depends on the array properties. The simplest spherical array configuration is the open sphere composed of sensors suspended in free space. It is assumed that the other accessories, such as cables and mounting brackets, have no effect for sensors capturing. However, this configuration might make the open sphere suffer from poor robustness at some certain frequency points. Another common configuration is the rigid sphere. In this configuration, sensors are mosaicked on a rigid spherical baffle, which means the sound waves will be scattered by the sphere. So, the mode strength of these two configurations is different. The magnitude of $b_n(kR)$ for different sphere is shown in Fig. 2.

The specific form can be expressed as

$$b_n(kR) = \begin{cases} 4\pi i^n j_n(kR) & \text{open sphere} \\ 4\pi i^n \left(j_n(kR) - \dfrac{j_n'(kR)}{h_n'(kR)} h_n(kR)\right) & \text{rigid sphere,} \end{cases}$$
$$(5)$$

$Y_n^m(\theta, \phi)$ defined as:

$$Y_n^m(\theta, \phi) = \sqrt{\frac{2n+1}{4\pi} \frac{(n-m)!}{(n+m)!}} P_n^m(\cos\theta) e^{im\phi}, \qquad (6)$$

where $0 \le n \le N$, $-n \le m \le n$, and $P_n^m(\cos\theta)$ are the associated Legendre polynomials [32].

According to (4), the dictionary matrix $\hat{\mathbf{A}}$ can be written as

$$\hat{\mathbf{A}} = \mathbf{Y}(\mathbf{\Omega})\mathbf{B}(k)\mathbf{Y}^H(\mathbf{\Phi}), \qquad (7)$$

where $\mathbf{Y}(\mathbf{\Omega})$ is an $L \times J$ spherical harmonic matrix whose lth row is given as:

$$\mathbf{y}(\Omega_l) = \Bigg[\underbrace{Y_0^0(\Omega_l)}_{n=0}, \underbrace{Y_1^{-1}(\Omega_l), Y_1^0(\Omega_l), Y_1^1(\Omega_l)}_{n=1} \qquad (8)$$
$$, \cdots, \underbrace{Y_N^{-N}(\Omega_l), \cdots, Y_N^N(\Omega_l)}_{n=N}\Bigg],$$

where $J = (N+1)^2$, $\mathbf{Y}(\mathbf{\Phi})$ is a $G \times J$ spherical harmonic matrix defined similarly to $\mathbf{Y}(\mathbf{\Omega})$, and $\mathbf{B}(k)$ is a $J \times J$ far-field mode strength matrix:

$$\mathbf{B}(k) = \text{diag}\Bigg(\underbrace{b_0(kR)}_{n=0}, \underbrace{b_1(kR), b_1(kR), b_1(kR)}_{n=1} \qquad (9)$$
$$, \cdots, \underbrace{b_N(kR), \cdots, b_N(kR)}_{n=N}\Bigg).$$

From (7), the dictionary matrix consists of three terms. The first and the second terms are only correlated with the sensor positions and the wavenumber, respectively. Furthermore, considering $\mathbf{Y}(\mathbf{\Omega})\mathbf{B}(k)$ is an $L \times J$ matrix ($L \ge J$) that has a left pseudo inverse, (3) can be rewritten as:

$$\overline{\mathbf{X}}_t = \overline{\mathbf{A}}\overline{\mathbf{S}}_t + \overline{\mathbf{V}}_t, \qquad (10)$$

where $\overline{\mathbf{X}}_t = \mathbf{B}^{-1}(k)\mathbf{Y}^+(\mathbf{\Omega})\breve{\mathbf{X}}_t$, $\overline{\mathbf{A}} = \mathbf{Y}^H(\mathbf{\Phi})$, and $\overline{\mathbf{V}}_t = \mathbf{B}^{-1}(k)\mathbf{Y}^+(\mathbf{\Omega})\breve{\mathbf{V}}_t$. Owing to the special property of spherical harmonic function, the complex-valued model can be transformed into a real-valued one to reduce the computational complexity.

Because of the characterization of $[Y_n^m(\Phi)]^* = (-1)^m Y_n^{-m}(\Phi)$ for each element in $\overline{\mathbf{A}}$, we can transform $\overline{\mathbf{A}}$ into a matrix with column conjugation property [33] for each order by $\mathbf{Q}_1$:

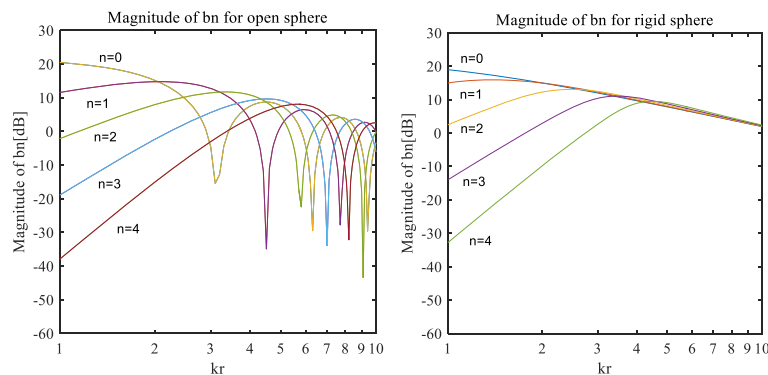**Fig. 2** The magnitude of bn for different sphere

$$\mathbf{Q}_1 \overline{\mathbf{X}}_t = \mathbf{Q}_1 \overline{\mathbf{A}} \mathbf{S}_t + \mathbf{Q}_1 \overline{\mathbf{V}}_t, \tag{11}$$

$$\mathbf{Q}_1 = \text{blkdiag}\left(\mathbf{Q}_{1,0}, \cdots, \mathbf{Q}_{1,N}\right), \tag{12}$$

$$\mathbf{Q}_{1,n} = \text{diag}\left(\underbrace{(-1)^n, (-1)^{n-1}\cdots, (-1)}_{n}, 1, \underbrace{1, 1, \cdots 1}_{n}\right). \tag{13}$$

$\mathbf{Q}_1$ is a unitary matrix. The new matrix $\mathbf{Q}_1\overline{\mathbf{A}}$ is:

$$\mathbf{Q}_1\overline{\mathbf{A}} = \left[\mathbf{y}_0^0(\boldsymbol{\Phi}), \cdots, \underbrace{\mathbf{y}_N^N(\boldsymbol{\Phi}), \cdots, \mathbf{y}_N^0(\boldsymbol{\Phi}), \cdots, \left[\mathbf{y}_N^N(\boldsymbol{\Phi})\right]^*}_{2N+1}\right]^T, \tag{14}$$

where $\mathbf{y}_n^m(\boldsymbol{\Phi}) = \left[Y_n^m(\Phi_1), \cdots, Y_n^m(\Phi_G)\right]^T$. From (14), we can find that each order has the column conjugation property, which means we can transform the complex-valued matrix into a real-valued one by linear combinations. So, the transform matrix is built as:

$$\mathbf{Q}_2 = \text{blkdiag}\left(\mathbf{Q}_{2,0}, \cdots, \mathbf{Q}_{2,N}\right), \tag{15}$$

$$\mathbf{Q}_{2,n} = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{I}_n & \mathbf{0}_n & \mathbf{J}_n \\ \mathbf{0}_n^T & \sqrt{2} & \mathbf{0}_n^T \\ -i\mathbf{J}_n & \mathbf{0}_n & i\mathbf{I}_n \end{bmatrix}, \tag{16}$$

where $n \geq 1$, and $\mathbf{Q}_{2,0} = 1$. From (15), $\mathbf{Q}_2$ is a block diagonal unitary matrix due to $\mathbf{Q}_2^{-1} = \mathbf{Q}_2^H$. Utilizing the unitary transform matrix $\mathbf{Q}_2$, (11) can be changed into

$$\mathbf{Q}_2\mathbf{Q}_1\overline{\mathbf{X}}_t = \mathbf{Q}_2\mathbf{Q}_1\overline{\mathbf{A}}\mathbf{S}_t + \mathbf{Q}_2\mathbf{Q}_1\overline{\mathbf{V}}_t. \tag{17}$$

We use $\tilde{\mathbf{X}}_t = \mathbf{A}\overline{\mathbf{S}}_t + \tilde{\mathbf{V}}_t$ to stand for (17) for simplicity. The new signal and noise matrices after the unitary transformation can be described as $\tilde{\mathbf{X}}_t = \mathbf{Q}\overline{\mathbf{X}}_t$ and $\tilde{\mathbf{V}}_t = \mathbf{Q}\overline{\mathbf{V}}_t$, where $\mathbf{Q} = \mathbf{Q}_2\mathbf{Q}_1$. The real-valued dictionary matrix can be inferred as follows:

$$\mathbf{A} = \mathbf{Q}\overline{\mathbf{A}} = \mathbf{Q}\mathbf{Y}^H(\boldsymbol{\Phi}) = \hat{\mathbf{Y}}^H(\boldsymbol{\Phi})$$
$$= \left[\hat{\mathbf{y}}_0^0(\boldsymbol{\Phi}), \cdots, \underbrace{\hat{\mathbf{y}}_N^{-N}(\boldsymbol{\Phi}), \cdots, \hat{\mathbf{y}}_N^0(\boldsymbol{\Phi}), \cdots, \hat{\mathbf{y}}_N^N(\boldsymbol{\Phi})}_{2N+1}\right]^T, \tag{18}$$

where $\hat{\mathbf{y}}_n^m(\boldsymbol{\Phi}) = \left[\hat{Y}_n^m(\Phi_1), \cdots, \hat{Y}_n^m(\Phi_G)\right]^T$ and $\hat{Y}_n^m(\Phi_g)$ is

$$Y_n^m(\Phi_g) = \begin{cases} \sqrt{2}\,\text{Re}\left(Y_n^m(\Phi_g)\right) & m < 0 \\ Y_n^m(\Phi_g) & m = 0 \\ (-1)^{m-1}\sqrt{2}\,\text{Im}\left(Y_n^m(\Phi_g)\right) & m > 0 \end{cases} \tag{19}$$

Afterwards, $\tilde{\mathbf{X}}_t$, $\overline{\mathbf{S}}_t$, and $\tilde{\mathbf{V}}_t$ can be partitioned into real and imaginary parts which are combined as $\mathbf{X}_t = [\text{Re}(\tilde{\mathbf{X}}_t), \text{Im}(\tilde{\mathbf{X}}_t)]$, $\mathbf{S}_t = [\text{Re}(\overline{\mathbf{S}}_t), \text{Im}(\overline{\mathbf{S}}_t)]$, and $\mathbf{V}_t = [\text{Re}(\tilde{\mathbf{V}}_t), \text{Im}(\tilde{\mathbf{V}}_t)]$. Therefore, the complex model (11) can be transformed into a real-valued one as:

$$\mathbf{X}_t = \mathbf{A}\mathbf{S}_t + \mathbf{V}_t. \tag{20}$$

Now, the real-valued signal model (20) is obtained in the spherical harmonics domain.

3 Variational sparse Bayesian learning embedded with Kalman filter (VSBLKF) for 2-D DOA tracking

When signal sources move in the whole angular space, it is necessary to model the evolution of $\mathbf{S}_t$ in time. Most of existing tracking state models [19, 20, 34] reveal the velocity information of the target movement. However, a TP model [35] is adopted to describe the movement of sources in a more practical way in this paper. It assumes that each source can transfer to its adjacent grids or stay at the current grid with an equal probability. So, the grid set $\mathbf{D} = \{1, \cdots, G\}$ distributes in a plane with G_1 rows and G_2 columns and is divided into three types, the corner area ($\mathbf{D}_1$, $\mathbf{D}_3$, $\mathbf{D}_7$, and $\mathbf{D}_9$), the marginal area ($\mathbf{D}_2$, $\mathbf{D}_4$, $\mathbf{D}_6$, and $\mathbf{D}_8$), and the center area ($\mathbf{D}_5$). The corresponding relation between the three areas and the angular grids (Φ_g) is shown in Fig. 3. $\mathbf{S}_{t-1}(g_l, :)$ denotes the vector at the g_lth grid during the time interval $t-1$. The probability of the vector moving to the g_cth position during the time interval t is f_{ij} as follows
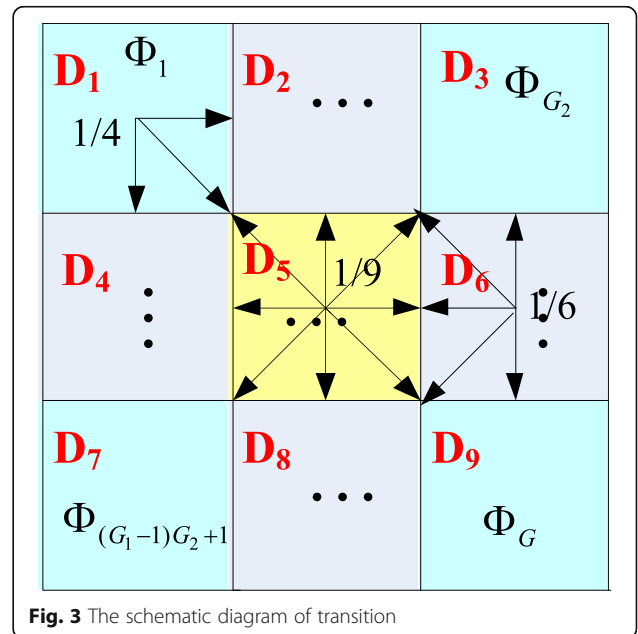

Fig. 3 The schematic diagram of transition

$$
f_{g_l g_c} = \begin{cases}
1/4 & g_l = 1, g_c - g_l = \{0, 1, G_2, G_2 + 1\} \\
1/6 & g_l \in \mathbf{D}_2, g_c - g_l = \{0, \pm 1, G_2, G_2 \pm 1\} \\
1/4 & g_l = G_2, g_c - g_l = \{0, -1, G_2, G_2 - 1\} \\
1/6 & g_l \in \mathbf{D}_4, g_c - g_l = \{0, 1, \pm G_2, 1 \pm G_2\} \\
1/9 & g_l \in \mathbf{D}_5, g_c - g_l = \{0, \pm 1, \pm(G_2 - 1), \pm(G_2 + 1), \pm G_2\} \\
1/6 & g_l \in \mathbf{D}_6, g_c - g_l = \{0, -1, \pm G_2, -1 \pm G_2\} \\
1/4 & g_l = (G_1 - 1)G_2 + 1, g_c - g_l = \{0, 1, -G_2, 1 - G_2\} \\
1/6 & g_l \in \mathbf{D}_8, g_c - g_l = \{0, \pm 1, -G_2, -G_2 \pm 1\} \\
1/4 & g_l = G, g_c - g_l = \{0, -1, -G_2, -G_2 - 1\} \\
0 & otherwise
\end{cases}
\tag{21}
$$

Therefore, the vector in the g_cth position at the time interval t can be expressed as $\mathbf{S}_t(g_c, :) = \sum_{g_l = 1}^{G} f_{g_l g_c} \mathbf{S}_{t=1}(g_l, :)$. Considering the state noise exists in the process of transition, the state transition model can be described as

$$
\mathbf{S}_t = \mathbf{F}\mathbf{S}_{t-1} + \mathbf{E}_t, \tag{22}
$$

where $\mathbf{E}_t$ is the state noise.

In tracking dynamic targets, the most popular method is KF technique. It is based on Gaussian assumption. In the standard KF method, the covariance matrix of the state noise and the variance of the measurement noise are assumed to be known. However, in real applications, it is difficult to know these parameters. Because of the advantage of VSBL, which can model the sparse signal well, we want to introduce VSBL into the KF method to estimate the covariance matrix of the state noise and the measurement noise variances. Nevertheless, based on the model we construct, it might be infeasible to calculate the posterior distribution. Because the dimensionality of the latent parameters is too high to calculate directly. Therefore, we attempt to turn to approximation schemes for help. The marginal probability of the observed data $\mathbf{X}$ is gotten by integrating over the remaining unobserved variables $\boldsymbol{\theta}$(including the latent variables $\mathbf{S}$ and some hyperparameters) [36].

$$
P(\mathbf{X}) = \int P(\mathbf{X}, \boldsymbol{\theta}) d\boldsymbol{\theta}. \tag{23}
$$

However, it is knotty to solve this integration. The variational approaches can be used to solve this problem by pulling in a distribution $Q(\boldsymbol{\theta})$ which allows the marginal log likelihood to be decomposed into two terms

$$
\ln P(\mathbf{X}) = L(Q) + KL(Q\|P), \tag{24}
$$

$$
L(Q) = \int Q(\boldsymbol{\theta}) \ln \frac{P(\mathbf{X}, \boldsymbol{\theta})}{Q(\boldsymbol{\theta})} d\boldsymbol{\theta}, \tag{25}
$$

where $KL(Q\|P)$ is the Kull-Leibler divergence between $Q(\boldsymbol{\theta})$ and the posterior distribution $P(\boldsymbol{\theta}|\mathbf{X})$ and it is given by

$$
KL(Q\|P) = -\int Q(\boldsymbol{\theta}) \ln \frac{P(\boldsymbol{\theta}|\mathbf{X})}{Q(\boldsymbol{\theta})} d\boldsymbol{\theta}, \tag{26}
$$

$L(Q)$ is a functional about $Q(\boldsymbol{\theta})$. Since $KL(Q\|P) \geq 0$, it follows that $L(Q)$ is strictly lower bound on $\ln P(\mathbf{X})$. Furthermore, because the left side of (24) is independent of $Q(\boldsymbol{\theta})$, maximizing $L(Q)$ is tantamount to minimizing $KL(Q\|P)$. Therefore $Q(\boldsymbol{\theta})$ represents an approximation to the posterior distribution $P(\boldsymbol{\theta}|\mathbf{X})$. The goal in a variational approach is to choose a suitable form for $Q(\boldsymbol{\theta})$ which is adequately simple and flexible. Besides, it should make the lower bound $L(Q)$ to be readily evaluated and tight. In fact, some family of $Q(\boldsymbol{\theta})$ distributions can be chosen and the best approximation within this family is found by maximizing the lower bound with respect to $Q(\boldsymbol{\theta})$. One approach is to assume some tractable forms for $Q(\boldsymbol{\theta})$ and then to optimize $L(Q)$ about the parameters of the distribution [36, 37]. We adopt an alternative step and consider a factorized form over the component variables $\{\theta_\zeta\}$ in $\boldsymbol{\theta}$ as follows

$$
Q(\boldsymbol{\theta}) = \prod_\zeta Q_\zeta(\theta_\zeta). \tag{27}
$$

This scheme is called a mean field theory [38]. Because of the true posterior distribution might not be calculated, an approximate distribution is introduced to estimate the distribution. In this paper, the mean field theory is adopted as the tool to solve above problem. In this method, we only limit the type range of probability distribution rather than the form of probability distribution, which can make the approximate distribution become more flexible and extensive. In the scenario considered, the signal is sparse in the whole space at each time interval, the potential position and noise variance can be seen as the unobserved variables, and the array receiving the signal can be seen as the observation variables. Therefore, the final model can be solved with the VSBL [39]. In the mean field theory of the posterior, the algorithm can optimize one parameter at a time while fixing all other parameters. The optimal distribution for each of the parameters can be expressed as:

$$
\ln Q(\boldsymbol{\mu}) \propto \left\langle \ln p\left(\mathbf{X}, \theta\right) \right\rangle_{\boldsymbol{\theta}\backslash\boldsymbol{\mu}}, \tag{28}
$$

where $\boldsymbol{\mu}$ indicates one of the parameter and $\langle \ln p(\mathbf{X}, \boldsymbol{\theta})\rangle_{\boldsymbol{\theta}\backslash\boldsymbol{\mu}}$ is the expectation of the joint probability of the data and latent variables which take over all variables except $\boldsymbol{\mu}$.

According to the Eq. (22), the state noise is assumed as the Gaussian distribution. The prior distribution of $\mathbf{E}_t$ is

Huang et al. EURASIP Journal on Advances in Signal Processing (2018) 2018:23

Page 7 of 14

$$p(\mathbf{E}_t) = \mathcal{MN}_{G,2B}\big(0, \mathbf{\Sigma}_{t|t}, \mathbf{I}_G\big)$$
$$= (2\pi)^{-G}|\mathbf{\Delta}_t|^{-B}\exp\left\{-\frac{1}{2}\mathrm{tr}\Big(\mathbf{\Sigma}_{t|t}^T\mathbf{\Delta}_t^{-1}\mathbf{\Sigma}_{t|t}\Big)\right\}.$$
(29)

Since the measurement noise satisfies Gaussian distribution, we can get

$$p(\mathbf{V}_t) = \mathcal{MN}_{J,2B}(0, \mathbf{\Delta}_t, \mathbf{I}_{2B})$$
$$= (2\pi)^{-J}|\mathbf{\Delta}_t|^{-B}\exp\left\{-\frac{1}{2}\mathrm{tr}\big(\mathbf{V}_t^T\mathbf{\Delta}_t^{-1}\mathbf{V}_t\big)\right\}, \quad (30)$$

where $\mathbf{\Delta}_t = \mathrm{diag}(\Delta_{1,\ t}, \cdots, \Delta_{J,\ t})$. The likelihood of measurement signal can be expressed as

$$p(\mathbf{X}_t|\mathbf{S}_t, \mathbf{\Delta}_t) = \mathcal{MN}_{J,2B}(\mathbf{AS}_t, \mathbf{\Delta}_t, \mathbf{I}_{2B})$$
$$= (2\pi)^{-J}|\mathbf{\Delta}_t|^{-B}\exp\left\{-\frac{1}{2}\mathrm{tr}\Big[(\mathbf{X}_t-\mathbf{AS}_t)^T\mathbf{\Delta}_t^{-1}(\mathbf{X}_t-\mathbf{AS}_t)\Big]\right\}.$$
(31)

Equivalently, we introduce a conjugate prior over the inverse noise variance $\mathbf{\Delta}_t$ given by Gamma distribution:

$$p\big(\mathbf{\Delta}_t^{-1}\big) = \prod_{j=1}^{J} Gam\Big(\Delta_{j,t}^{-1}|\mathbf{b}_{j,t}^{(\Delta)}, \mathbf{c}_{j,t}\Big). \quad (32)$$

The Eq. (31) will be the exponential form of quadric function about variable $\mathbf{S}_t$, when the variable $\mathbf{\Delta}_t$ is seen as a constant. Therefore, the conjugate prior of $\mathbf{S}_t$ will obey the Gaussian distribution [36]. So, the prior distribution of $\mathbf{S}_t$ is

$$p(\mathbf{S}_t|\mathbf{FS}_{t-1}, \mathbf{\Lambda}_t) = \mathcal{MN}_{G,2B}(\mathbf{FS}_{t-1}, \mathbf{\Lambda}_t, \mathbf{I}_G)$$
$$= (2\pi)^{-G}|\mathbf{\Lambda}_t|^{-B}\exp\left\{-\frac{1}{2}\mathrm{tr}\Big[(\mathbf{S}_t-\mathbf{FS}_{t-1})^T\mathbf{\Lambda}_t^{-1}(\mathbf{S}_t-\mathbf{FS}_{t-1})\Big]\right\}.$$
(33)

The signal satisfies Gaussian distribution with zero mean. The variance of $\mathbf{S}_t$ is $\mathbf{\Lambda}_t = \mathrm{diag}(\boldsymbol{\alpha}_t^{-1})$, where hyperparameter $\boldsymbol{\alpha}_t = [\alpha_{1,\ t}, \cdots, \alpha_{G,\ t}]$. At the same time, the conjugate prior of $\alpha_{g,\ t}$ satisfies Gamma distribution

$$p(\alpha_{g,t}) = Gam(\alpha_{g,t}|a_{g,t}, b_{g,t}) = b_{g,t}^{a_{g,t}}\alpha_{g,t}^{a_{g,t}-1}e^{-b_{g,t}\alpha_{g,t}}/\Gamma(a_{g,t}),$$
(34)

where $\Gamma(\cdot)$ is a Gamma function [40]. The expectation of $\alpha_{g,\ t}$ in (34) is $\langle\alpha_{g,\ t}\rangle = a_{g,\ t}/b_{g,\ t}$. Note that the marginal distribution of $\mathbf{S}_t$ can be obtained by integrating over $\boldsymbol{\alpha}_t$. Note that if the prior distribution and likelihood function conjugate each other, it will make the posterior distribution to have the same form with prior distribution [36].

Using the chain rule of probability [39], the posterior probabilistic distribution can be expressed as:

$$p(\boldsymbol{\alpha}_t, \mathbf{S}_t, \mathbf{\Delta}_t|\mathbf{X}_t)\propto p(\mathbf{X}_t|\mathbf{S}_t, \mathbf{\Delta}_t)p(\mathbf{S}_t|\boldsymbol{\alpha}_t)p(\boldsymbol{\alpha}_t)p(\mathbf{\Delta}_t).$$
(35)

The variational framework introduces a factorial representation (27) approximating to the posterior distribution $p(\boldsymbol{\alpha}_t, \mathbf{S}_t, \mathbf{\Delta}_t|\mathbf{X}_t)$ by $Q(\boldsymbol{\alpha}_t, \mathbf{S}_t, \mathbf{\Delta}_t) = Q(\mathbf{S}_t)Q(\boldsymbol{\alpha}_t)$ $Q(\mathbf{\Delta}_t)$. We propose the VSBLKF algorithm using the expectation maximization (EM) updates. In order to compute the E-step, it needs to know the posterior distribution of the unknown sparse state signal, which is Gaussian with mean $\mathbf{U}_{t|t}$ and covariance $\mathbf{\Sigma}_{t|t}$. These parameters can be computed by using KF prediction and updating equations as follows:

Predict steps:

$$\mathbf{U}_{t|t-1} = \mathbf{F}\mathbf{U}_{t-1|t-1}, \quad (36)$$

$$\mathbf{\Sigma}_{t|t-1} = \mathbf{F}\mathbf{\Sigma}_{t-1|t-1}\mathbf{F}^T + \langle\mathbf{\Lambda}_t\rangle_q. \quad (37)$$

Update steps:

$$\mathbf{K}_t = \mathbf{\Sigma}_{t|t-1}\mathbf{A}^T\Big(\mathbf{A}\mathbf{\Sigma}_{t|t-1}\mathbf{A}^T + \langle\mathbf{\Delta}\rangle_q\Big)^{-1}, \quad (38)$$

$$\mathbf{U}_{t|t} = \mathbf{U}_{t|t-1} + \mathbf{K}_t\big(\mathbf{X}_t-\mathbf{A}\mathbf{U}_{t|t-1}\big), \quad (39)$$

$$\mathbf{\Sigma}_{t|t} = \mathbf{\Sigma}_{t|t-1}-\mathbf{K}_t\mathbf{A}\mathbf{\Sigma}_{t|t-1}. \quad (40)$$

Here, the initial value $\mathbf{U}_{0|0}$ can be assumed to be known. In the M-step, we use VSBL to calculate the state variance and measurement variance of the current time. According to the Eq. (28)

$$\ln Q(\boldsymbol{\alpha}_t)\propto\Big\langle\ln p(\mathbf{S}_t|\mathbf{FS}_{t-1}, \mathbf{\Lambda}_t)p(\boldsymbol{\alpha}_t)\Big\rangle_{\theta\backslash\alpha_t}$$
$$= \Big\langle\ln p(\mathbf{S}_t|\mathbf{FS}_{t-1}, \mathbf{\Lambda}_t)\Big\rangle + \ln p(\boldsymbol{\alpha}_t)$$
$$= -B\ln|\mathbf{\Lambda}_t|-\frac{1}{2}tr\Big\{\Big\langle(\mathbf{S}_t-\mathbf{FS}_{t-1})^T\mathbf{\Lambda}_t^{-1}(\mathbf{S}_t-\mathbf{FS}_{t-1})\Big\rangle\Big\}$$
$$+ \sum_{g=1}^{G}\Big\{(a_{g,t}-1)\ln\alpha_{g,t}-b_{g,t}\alpha_{g,t}\Big\} + C$$
$$= \sum_{g=1}^{G}\Big\{(a_{g,t}-1+B)\ln\alpha_{g,t}-\Big(b_{g,t}+\frac{1}{2}\Big\langle(\mathbf{S}_t-\mathbf{FS}_{t-1})(\mathbf{S}_t-\mathbf{FS}_{t-1})^T\Big\rangle\Big)\alpha_{g,t}\Big\} + C$$
(41)

C is a constant number. Therefore, the posterior is given as:

$$Q(\boldsymbol{\alpha}_t) = \prod_{g=1}^{G} Gam\big(\alpha_{g,t}|\overline{a}_{g,t}, \overline{b}_{g,t}\big), \quad (42)$$

$$\overline{a}_{g,t} = a_{g,t} + B, \quad (43)$$

$$\overline{b}_{g,t} = b_{g,t} + 0.5\Big\langle(\mathbf{S}_t-\mathbf{FS}_{t-1})(\mathbf{S}_t-\mathbf{FS}_{t-1})^T\Big\rangle, \quad (44)$$

$$\langle \Lambda_{g,t} \rangle \ = \overline{a}_{g,t}/\overline{b}_{g,t}, \tag{45}$$

The posterior distribution of the noise inverse variance can be similarly calculated as

$$Q(\mathbf{\Delta}_t) = \prod_{j=1}^{J} Gam\left(\Delta_{j,t} | \overline{b}_{j,t}^{(\Delta)}, \overline{c}_{j,t}\right). \tag{46}$$

$$\overline{c}_{j,t}^{(\Delta)} = c_{j,t}^{(\Delta)} + 0.5 \left\langle (\mathbf{X}_t - \mathbf{A}\mathbf{S}_t)(\mathbf{X}_t - \mathbf{A}\mathbf{S}_t)^T \right\rangle, \tag{47}$$

$$\overline{b}_{j,t}^{(\Delta)} = b_{j,t}^{(\Delta)} + J, \tag{48}$$

$$\langle \Delta_{j,t} \rangle \ = \overline{b}_{j,t}^{(\Delta)}/\overline{c}_{j,t}, \tag{49}$$

where the hyperpriors $a_{g,t}, b_{g,t}, b_{j,t}^{(\Delta)}, c_{j,t}$ are set as 10^{-3}. From (38), we can see that using these recursions cannot guarantee the prediction sparse. Running (38) up to a large enough t, the signal strength of this target will spill over to all entries of $\mathbf{U}_{t \mid t}$. So, we propose a state corrector method to determine the true moving direction of next time and guarantee the prediction sparse. For each signal, there are different possible positions at next time. These possible positions correspond to nonzero values in $\mathbf{U}_{t \mid t}$ noted as $\mathbf{U}_{t|t}(p_1, :), \cdots, \mathbf{U}_{t|t}(p_{\sum_{d=1}^{D} w_d}, :)$, where $w_d \in \{4, 6, 9\}$. The value in the set correspond to different signal positions—corner area, marginal area, and central area successively. We find D maximum values in $\mathbf{U}_{t \mid t}$ corresponding to the maximum of $\|[\mathbf{A}(:,p_\kappa)]^T \mathbf{X}_t\|_2$, where $\kappa = 1, \cdots, \sum_{d=1}^{D} w_d$, and consider these positions as true directions. At the same time, the rest of the rows in $\mathbf{U}_{t \mid t}$ are set to zeros. The steps of the proposed algorithm are summarized in the following. After that, we discuss the computational complexity of two models using VSBLKF to track DOA trajectory. The two models are the traditional complex-valued model and the proposed real-valued model, respectively. Different models lead to different measurement matrix dimensions, which may affect the parameter updates in VSBL. Table 1 shows the amount of computations using the two models. $\mathbf{A}_{t,\ c}, \mathbf{X}_{t,\ c}, \mathbf{S}_{t,\ c}$ and $\mathbf{A}_{t,\ r}, \mathbf{X}_{t,\ r}$ and $\mathbf{S}_{t,\ r}$ represent the measurement matrices,

Table 1 Comparison of computational load in one iteration

		Real multiplication	Real addition
Complex-valued model	$\mathbf{A}_{t,c}^H \mathbf{\Sigma}_{t|t-1}$	$4JG^2$	$(4J-2)G^2$
	$\mathbf{A}_{t,c}^H \mathbf{S}_{t,c}$	$4JG$	$(4G-2)J$
Real-valued model	$\mathbf{A}_{t,r}^H \mathbf{\Sigma}_{t|t-1}$	$2JG^2$	$2(J-1)G^2$
	$\mathbf{A}_{t,r}^H \mathbf{S}_{t,r}$	$2JG$	$2(G-1)J$

receiving data and sparse signals in the complex-valued model and real-valued model. Real multiplication and real addition are used to evaluate the computational cost. Our proposed real model transforms the complex-valued measurement matrix into a real-valued one through a unitary transformation without changing the dimension. After the real-valued transform, the computation complexity for each iteration will reduce much when the steering matrix is used to calculate. Therefore, using the real-valued transform can improve the rate of the proposed algorithm. A unitary transformation needs $4J^2T$ real multiplications and $(4J-2)JT$ real additions, and it does not join the iteration of SBL. The computational load of the real-valued model is much lower than that of the complex-valued model, so the computation complexity can be reduced greatly.

Algorithm 1 VSBLKF algorithm

INPUTS:

1) The prior information for noise variance.

2) Initialize the prior parameters in (32), (34), (22) and (23).

3) Get the model in the spherical harmonics domain using (10).

Initial Configuration:

1: Construct the unitary matrix $\mathbf{Q}$ to transform the complex spherical harmonic signal into real spherical harmonic signal using (12), (15), (20), and (21);

2: Select a suitable value for convergence Thr.

3: Select a start value for noise variance and mean.

Repeat

4: Compute the prediction part of KF to coarsely estimate the latent variable means in (36).

Repeat

5: Use the rest parts of KF to estimate the latent variable means and variances in (37)-(40).

6: Use the variational Bayesian to evaluate the state variances and measurement variance of the current time in (42)-(49).

7: **Until** $\left\| \mathbf{U}_{t_t} - \mathbf{U}_{t|t}^- \right\|_2 < Thr$. ($\mathbf{U}_{q_t}^-$ is the estimation from the previous iteration)

8: Use the state corrector method to determine the true moving direction of next time and guarantee the prediction sparse.

9: **Until** $t = T$.

OUTPUTS: Return the Estimated positions — the non-zeros position of $\mathbf{U}_{t_t}$ ($t = 1, \cdots, T$).

4 Extension to off-grid problem
The assumption that the estimated signal directions lie on specified grids is unrealistic. To reduce the

Huang *et al. EURASIP Journal on Advances in Signal Processing* (2018) 2018:23

Page 9 of 14

errors which are caused by mismatch of the real locations and partition locations, the denser grids can be applied when dividing the measurement matrix of signal. Nevertheless, each column of the measurement matrix needs to satisfy the cross-unrelated property for sparse recovery. The denser grids might lead to a higher complexity and make the dictionary become coherent that violates the necessary conditions for compressive sensing. Hence, to make the proposed algorithm more practical, the off-grid model is considered. As mentioned earlier, the directional matrix $\hat{\mathbf{Y}}^H(\boldsymbol{\Phi})$ in the VSBLKF algorithm is constructed on the uniform sampled angle grids $\boldsymbol{\Phi} = \{\Phi_1, \cdots, \Phi_G\}$ with $\Phi_g = (\theta_g, \phi_g)$, $g = 1, \cdots, G$. The target signal angles are $\breve{\boldsymbol{\Phi}} = \{\breve{\Phi}_1, \cdots, \breve{\Phi}_D\}$ with $\breve{\Phi}_d = (\breve{\theta}_d, \breve{\phi}_d)$, $d = 1, \cdots, D$, D<<G, and $\tilde{\Phi}_d$ which is not equal to Φ_g. In addition, we define the nearest grid of target signal angles as $\hat{\Phi}_{g_d} = (\hat{\theta}_{g_d}, \hat{\phi}_{g_d})$. The true steering vector $\mathbf{a}(\tilde{\Phi}_d)$ can be approximated by its Taylor extension as [41].

$$\mathbf{a}(\tilde{\Phi}_d) \approx \mathbf{a}(\hat{\Phi}_{g_d}) + \mathbf{b}(\hat{\theta}_{g_d})(\breve{\theta}_d - \hat{\theta}_{g_d}) + \mathbf{c}(\hat{\phi}_{g_d})(\breve{\phi}_d - \hat{\phi}_{g_d}),$$
(50)

where $\mathbf{a}(\hat{\Phi}_{g_d})$ represents the original steering vector, $\mathbf{b}(\hat{\theta}_{g_d})$ and $\mathbf{c}(\hat{\phi}_{g_d})$ are the partial derivative of $\breve{\theta}_d$ and $\breve{\phi}_d$ with respect to $\hat{\Phi}_{g_d}$. Let

$$\mathbf{B} = [\mathbf{b}(\theta_1), \cdots, \mathbf{b}(\theta_G))],$$
(51)

$$\mathbf{C} = [\mathbf{c}(\phi_1), \cdots, \mathbf{c}(\phi_G)],$$
(52)

$$\boldsymbol{\beta}^{\mathrm{T}} = [\beta_1, \beta_2, \cdots, \beta_G],$$
(53)

$$\boldsymbol{\gamma}^{\mathrm{T}} = [\gamma_1, \gamma_2, \cdots, \gamma_G],$$
(54)

$$\beta_g = \begin{cases} \breve{\theta}_d - \hat{\theta}_{g_d}, & g = g_d \\ 0, & otherwise \end{cases},$$
(55)

$$\gamma_g = \begin{cases} \breve{\phi}_d - \hat{\phi}_{g_d}, & g = g_d \\ 0, & otherwise \end{cases}.$$
(56)

So, the new steering matrix $\mathbf{A} + \mathbf{B} \operatorname{diag}(\boldsymbol{\beta}) + \mathbf{C} \operatorname{diag}(\boldsymbol{\gamma})$ can replace the original steering matrix $\mathbf{A}$ to estimate both signal coarse girds and their bias with elevation and azimuth, simultaneously. According to [41], the author solved this problem from the Bayesian perspective. In this section, we use the least squares estimation to calculate the biases of elevation and azimuth, respectively. In the first step, we use VSBL to get the coarse estimation of true location. Next, using the expectation $\mathbf{U}_{t \mid t}$ to minimize the following equation as

$$\min_{\boldsymbol{\beta}^\xi, \boldsymbol{\gamma}^\xi} \|\boldsymbol{\beta}^\xi\|^2 + \|\boldsymbol{\gamma}^\xi\|^2 + \|\mathbf{X}_t - (\mathbf{A} + \mathbf{B} \operatorname{diag}(\boldsymbol{\beta}^\xi) + \mathbf{C} \operatorname{diag}(\boldsymbol{\gamma}^\xi))\mathbf{U}_{t|t}\|_F^2,$$
(57)

where $\boldsymbol{\beta}^\xi$ stands for the kth iteration update of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}^\xi$ represents the ξth iteration update of $\boldsymbol{\gamma}$. When updating difference biases, we keep another one fixed. It is tantamount to the least squares problem when we want to solve

$$\mathbf{X}_t - (\mathbf{A} + \mathbf{C}_t \operatorname{diag}(\boldsymbol{\gamma}^\xi))\mathbf{U}_{t|t}^\xi = \mathbf{B}\Xi_{\mathbf{S}_{t|t}}^\xi \boldsymbol{\beta}^\xi,$$
(58)

$$\mathbf{X}_t - (\mathbf{A} + \mathbf{B}_t \operatorname{diag}(\boldsymbol{\beta}^\xi))\mathbf{U}_{t|t}^\xi = \mathbf{C}\Xi_{\mathbf{S}_{t|t}}^\xi \boldsymbol{\gamma}^\xi,$$
(59)

Algorithm 2 OGVSBLKF algorithm

INPUTS:

1) The prior information for noise variance.

2) Initialize the prior parameters in (32), (34), (22) and (23).

3) Obtain the model in spherical harmonic domain using (10).

Initial Configuration:

1: Construct the unitary matrix $\mathbf{Q}$ to transform the complex spherical harmonic signal into real spherical harmonic signal using (12), (15), (20), and (21);

2: Select a suitable value for convergence Thr.

3: Select a start value for noise variance and mean.

Repeat

4: Compute the prediction part of KF to coarsely estimate the latent variable means in (36).

Repeat

5: Use the rest parts of KF to estimate the latent variable means and variances in (37)-(40).

6: Use the Variational Bayesian to evaluate the state variances and measurement variances of the current time in (42)-(49).

7: Calculate the off-grid biases using (60) and (61);

8: **Until** $\|\mathbf{U}_{t|t} - \mathbf{U}_{t|t}^-\|_2 < Thr$.

9: Use the state corrector method to determine the true moving direction of next time and guarantee the prediction sparse.

10: **Until** $t = T$.

OUTPUTS: Return the Estimated positions — the non-zero position of $\mathbf{U}_{t|t}$ ($t = 1, \cdots, T$).

---

where $\Xi_{\mathbf{S}_{t|t}}^\xi = \operatorname{diag}(S_{1,t|t}^\xi, \cdots, S_{G,t|t}^\xi)$. So, at the kth iteration update of $\boldsymbol{\beta}$ is:

$$\boldsymbol{\beta}^\xi = (\mathbf{B}\Xi_{\mathbf{S}_{t|t}}^\xi)^+ (\mathbf{X}_t - (\mathbf{A} + \mathbf{C} \operatorname{diag}(\boldsymbol{\gamma}^\xi))\mathbf{U}_{t|t}^\xi).$$
(60)

In a similar manner, the iterative form of $\boldsymbol{\gamma}$ can be obtained as:

$$\gamma^{\xi} = \left(\mathbf{C}\Xi^{\xi}_{\mathbf{S}_{t|t}}\right)^{+}\left(\mathbf{X}_t - \left(\mathbf{A} + \mathbf{B}\,\mathrm{diag}(\boldsymbol{\beta}^{\xi})\right)\mathbf{U}^{\xi}_{t|t}\right). \quad (61)$$

The initial value of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are set as $\mathbf{0}_G$ (a column vector containing G zeros). The number of source signals is known a priori. Each signal current coarse location can be approximated by VSBLKF and the biases between the reference and estimation are gauged using (59) and (60), respectively. The concrete steps are summarized in Algorithm 2.

5 Results and discussion

In this section, we want to verify the robustness and performance of the proposed algorithm compared with the standard KF, SBL, and VSBL methods. We use a rigid spherical array, which has 32 sensors distributed in a uniform way and radius $R = 0.1$ m. The maximum order of the spherical harmonics is $N = 4$. In our experiments, the ranges of elevation and azimuth are defined from 0° to 180° and from 0° to 360°. We divide them into 31 and 62 grids with stationary angular interval respectively. Therefore, there are 1922 grids which are the possible angles for source signals. Note that we only choose the range of azimuth from 0° to 180° in our simulations to reduce the calculate complexity. In the moving process, assuming each source is likely to move to its adjacent grids or stay at current grid, which means a source can move to the different directions with 6° or be static with equal probability. The proposed method can track other trajectories as long as the trajectory can be described by the grids and obey the TP model. The trajectory used in this paper is randomly generated under the frame of TP model. In order to show a series of performance quantitatively, such as on-grid, off-grid, and RMSE vs SNR, we used one trajectory to explain these results. One random realization of this movement model is considered for $T = 50$ time interval. The number of Monte Carlo trials is 500. The hyperpriors $a_{g,t}, b_{g,t}, b^{(\Delta)}_{j,t}, c_{j,t}$ are set as 10^{-3}.

5.1 Example 1: the performance of the proposed method with an on-grid model

In the first simulation, we show the tracking performance of one signal at each time interval. The signal to noise ratio (SNR) is set as 12 dB while the strength of the signal is 10. The moving trace obtained by the four methods are shown in Fig. 4. Figures 5 and 6 give the tracking performance at different time intervals. The performance of SBL is poor, because the likelihood function of the measurements involved in SBL cannot match the true one.
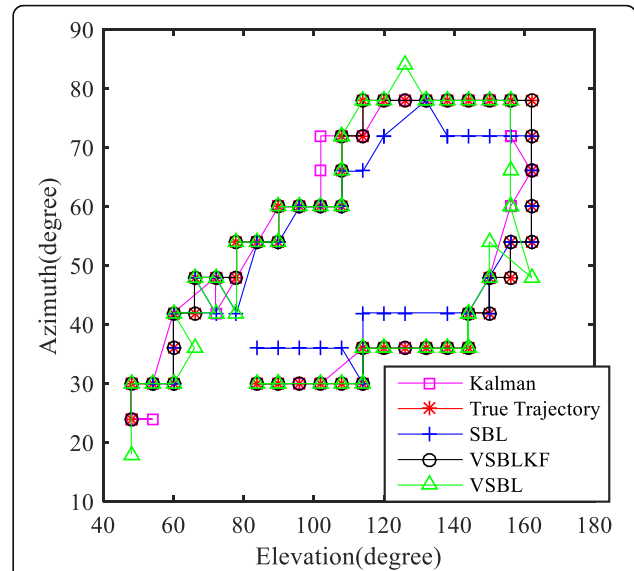

Fig. 4 Traces of four methods based in the on-grid model

We can see that the VSBLKF method achieves the best tracking performance because it combines the advantages of KF with VSBL method. Figures 7 and 8 denote the signal angle errors (the differences between the estimated angle and reference angle) varying with time intervals. From these curves, the proposed method is better than the other approaches.

5.2 Example 2: the RMSE versus SNR with an on-grid model

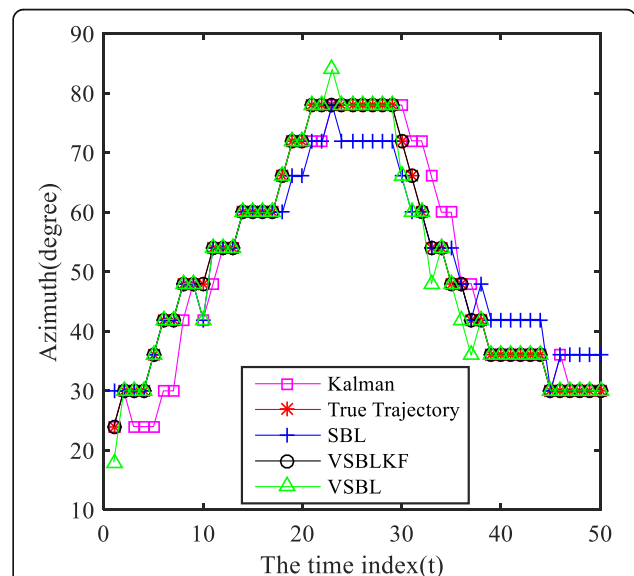In Fig. 9, we compare the performance of the proposed algorithm under different SNRs with other
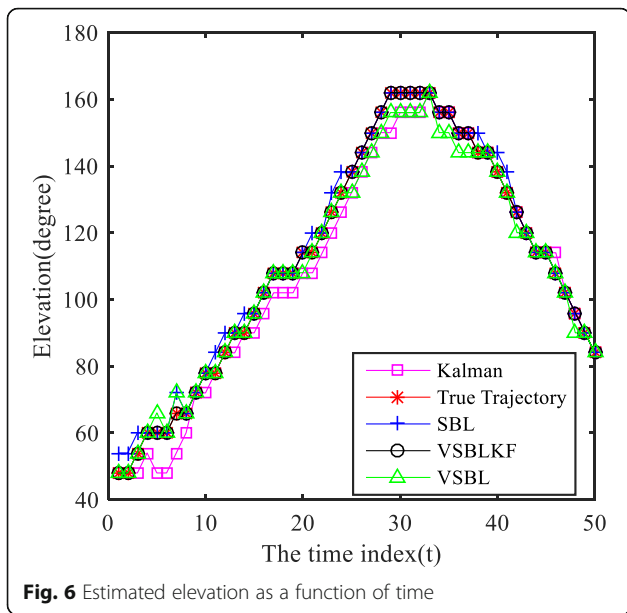

Fig. 5 Estimated azimuth as a function of time

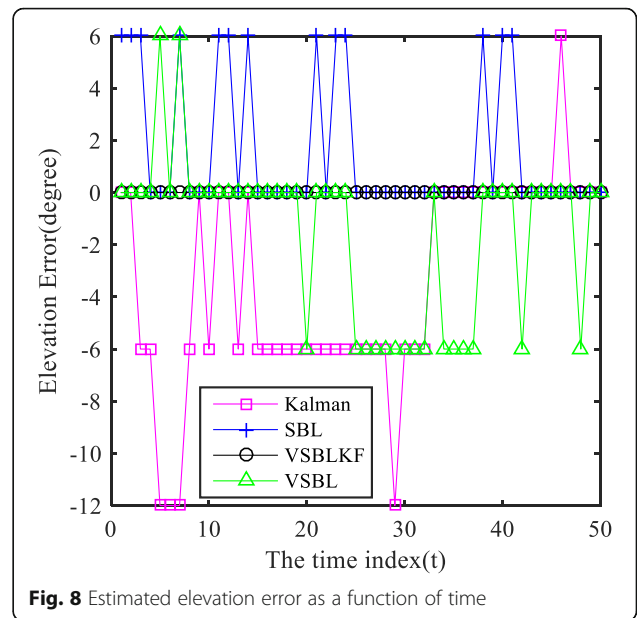**Fig. 6** Estimated elevation as a function of time

Fig. 8 Estimated elevation error as a function of time

approaches. The SNR is set as [7, 10, 13, 16, 20]. Root mean square error (RMSE) is adopted to measure the estimation performance under different SNRs. The RMSE is given by:

$$RMSE = \frac{1}{NUM} \sum_{n=1}^{NUM} \sqrt{\sum_{d=1}^{D} \sum_{t=1}^{T} \left(\breve{\theta}_{d,t} - \hat{\theta}_{d,t} \right)^2 + \left(\breve{\phi}_{d,t} - \hat{\phi}_{d,t} \right)^2 / 2DT},$$

(62)

where $\left(\breve{\theta}_{d,t}, \breve{\phi}_{d,t} \right)$ denotes the actual DOA, $\left(\hat{\theta}_{d,t}, \hat{\phi}_{d,t} \right)$ denotes the estimated DOA, *NUM* is the number of Monte Carlo trials, T is the time interval number, and

D is the signal numbers. We see that the VSBL method outperforms the KF method because the latter one does not consider the changes of the state noise variance. The prior knowledge of the variances of the state noise and the measured noise must be known for the KF approach. The proposed VSBLKF approach shows better performance than VSBL in tracking DOAs because the former one exploits the correlation of different time intervals. Just like the experiment display in Fig. 3, there are always some points deviating the ideal locations randomly.

5.3 Example 3: the performance of the proposed method with an off-grid model

In this part, we assume that the moving signal angular trajectory deviates the sampling grids. The deviating errors are set to be a standard normal distribution. The orbits of VSBLKF and OGVSBLKF are shown in Fig. 10, where the SNR is 12 dB and the time interval number is 50. Figure 10 shows that VSBLKF still cannot track the true locus reliably, but the proposed method can estimate the angular locus more accurately. Note that the OGVSBL method is not considered here because the OGVSBL method is an improving algorithm under the condition of locating in coarse grids correctly. As we can see in Fig. 4, the VSBL method cannot satisfy the demand yet. In Figs. 11 and 12, we can observe the specific errors between the estimated values and references clearly in the azimuth and elevation, respectively. We used the least-square estimation to calculate the deviation. However, we can find that it still has some deviation.
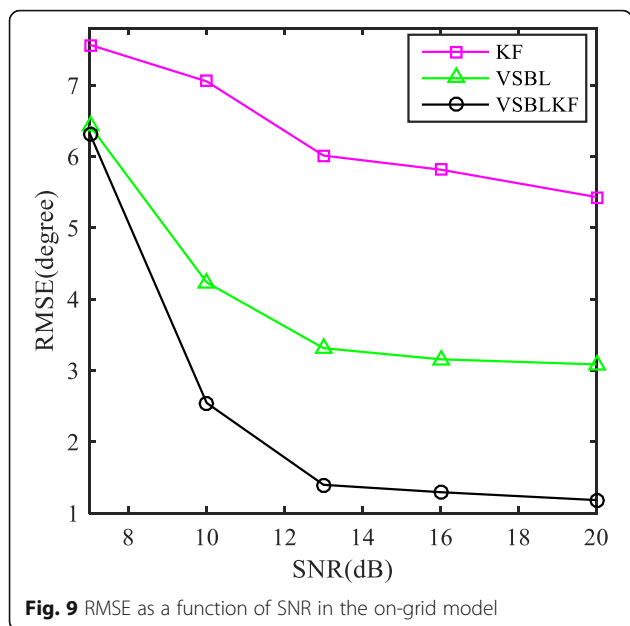

Fig. 7 Azimuth error as a function of time

Fig. 9 RMSE as a function of SNR in the on-grid model

Fig. 11 Estimated elevation as a function of time

It might be because the least square estimation we used still cannot handle the deviation problem well.

5.4 Example 4: the RMSE versus SNR with an off-grid model

In Fig. 13, we compare the performance of OGVSBLKF and VSBLKF under different SNRs varying from 7 to 20 dB with 3 dB interval approximately. In this simulation, we set the time interval number as 50 and Monte Carlo trials as 300. The angular deviating errors are set to be a normal distribution with a

variance of 1.5. The OGVSBLKF shows better performances than the VSBLKF method.

5.5 Example 5: the RMSE versus different grids

In Fig. 14, the influence of the proposed method with different grid intervals is depicted. From the figure, the estimation error will increase with the larger grid interval. Therefore, it is importance to select the suitable grid interval to estimate the signal direction trajectory when adopting the proposed method. However, this suitable grid interval might require several repeated tests. If we want to get a trade-off between performance and model complexity, we can use a coarse grid to search firstly, a
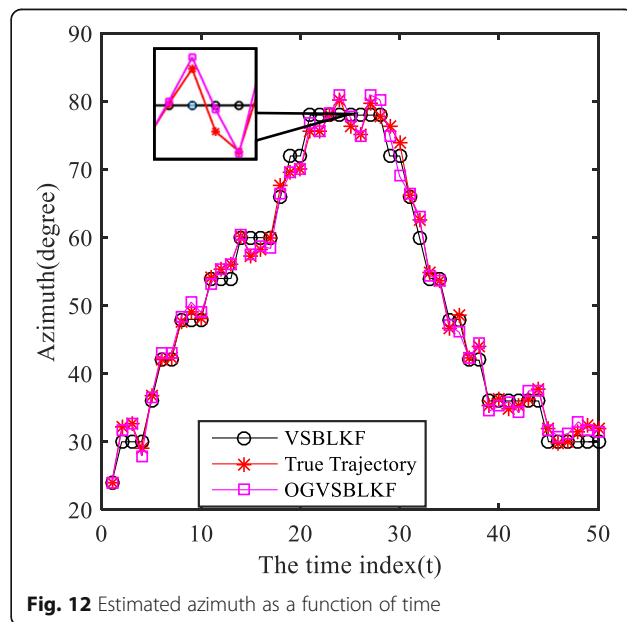

Fig. 10 Traces of two proposed methods in the off-grid model

Fig. 12 Estimated azimuth as a function of time

Huang *et al. EURASIP Journal on Advances in Signal Processing* (2018) 2018:23
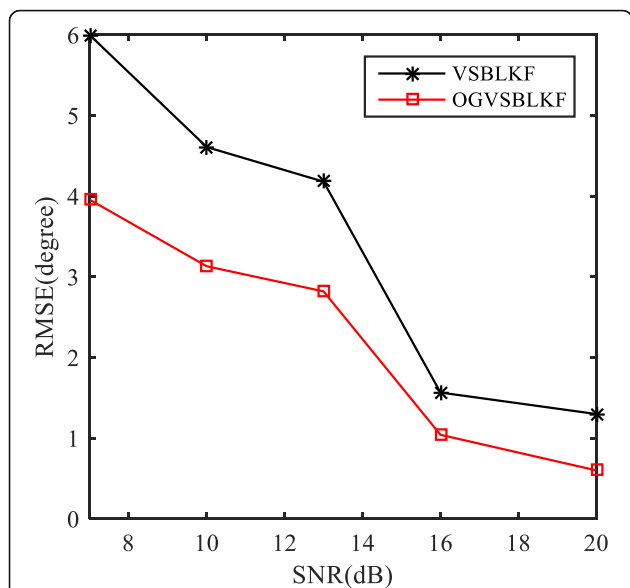
Page 13 of 14

Fig. 13 RMSE as a function of SNR in the off-grid model

fine grid then used to improve the precision of estimation.

5.6 Example 6: the performance of the proposed method for multiple signals

In this example, we compare our proposed algorithm with other methods in the condition of multiple moving signals impinging on the spherical array. It aims at illustrating that our proposed algorithm can be applicable to tracking multiple signals. Figure 15 is used to demonstrate the tracking performance when the
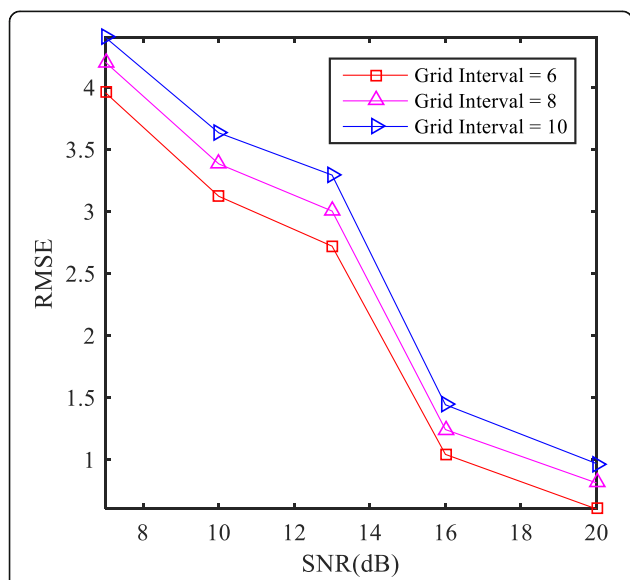

Fig. 14 RMSE as a function of SNR in the off-grid model with different grid interval
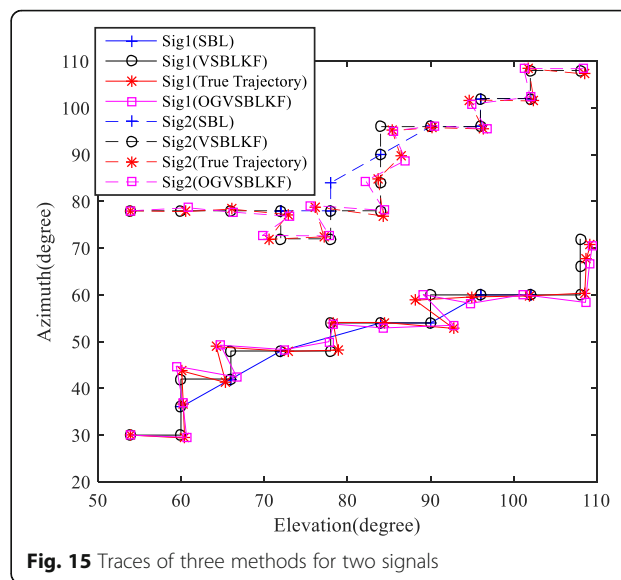

Fig. 15 Traces of three methods for two signals

SNR is 12 dB and the number of time intervals is 17 from the perspectives of elevation and azimuth. As we can see, SBL cannot track the signal trajectory reasonably and our proposed algorithm shows advantages in tracing multiple signals.

5.7 Example 7: the cost time versus different methods

In addition, the computational time of different DOA tracking methods is also analyzed. We conduct an evaluation of the computational complexity using TIC and TOC instruction in MATLAB. All the simulation results are obtained using the same PC with an Intel i7-6700 and 8 GB RAM, running MATLAB 2015b on 64-bit Windows 10. The average computational time is given in Table 2, which is obtained from 300 Monte Carlo trials and 50 time intervals. It can be seen that the time cost of the proposed method is more than that of VSBL, but the precision of the proposed method is higher. Comparing VSBLKF with OGVSBLKF, we observe that the method based on the off-grid model only needs 7 s more than that based on the on-grid model, so the performance of OGVSBLKF is better. Note that the theoretic complexity is $O(Tv)$ approximately, v is the iteration numbers. The reason of long cost time of the propose method is that using the Kalman filter embedded into the VSBL might increase the iteration numbers to avoid the local values.

Table 2 Comparison of time cost for different methods

	VSBL	VSBLKF (real)	OGVSBLKF (real)	VSBLKF (complex)	OGVSBLKF (complex)
Time(s)	4.51	21.23	28.51	44.31	55.36

6 Conclusions

In order to track the 2-D DOAs, we construct the state transition function according to the TP model based on a spherical array. The angular space is divided into grids to model a sparse signal. Through combining VSBL and KF methods, we propose an effective method called VSBLKF to track dynamic DOAs, where the KF estimation is embedded into the VSBL to estimate the signals. Besides, we extend our algorithm to the off-grid model. Simulations show that the proposed method achieves better tracking and anti-noise performance than VSBL and KF. In the future, we will extend the algorithm to wideband signals.

Authors' contributions
QH and JH designed and implemented the proposed algorithm and wrote the paper. KL and YF scientifically supervised the work and contributed in implementing the proposed algorithm. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References
1. A Hassanien, SA Vorobyov, Transmit energy focusing for DOA estimation in MIMO radar with colocated antennas. IEEE Trans. Signal Process. **59**(6), 2669–2682 (2011)
2. H Krim, M Viberg, Two decades of array signal processing research: the parametric approach. IEEE Signal Process. Mag. **13**(4), 67–94 (1996)
3. A Boukerche, H Oliveira, E Nakamura, A Loureiro, Localization systems for wireless sensor networks. IEEE Trans. Wirel. Commun. **14**(6), 6–12 (2007)
4. M Carlin, P Rocca, G Oliveri, Directions-of-arrival estimation through Bayesian compressive sensing strategies. IEEE Trans. Antennas Propag. **61**(7), 3828–3838 (2013)
5. B Wang, J Liu, X Sun, Mixed sources localization based on sparse signal reconstruction. IEEE Signal Process. Lett. **19**(8), 487–490 (2012)
6. D Malioutov, M Cetin, A Willsky, A sparse signal reconstruction perspective for source localization with sensor arrays. IEEE Trans. Signal Process. **53**(8), 3010–3022 (2005)
7. D Wipf, B Rao, An empirical Bayesian strategy for solving the simultaneous sparse approximation problem. IEEE Trans. Signal Process. **55**(7), 3704–3716 (2007)
8. S Ji, Y Xue, L Carin, Bayesian compressive sensing. IEEE Trans. Image Process. **56**(6), 2346–2356 (2008)
9. S Babacan, R Molina, A Katsaggelos, Bayesian compressive sensing using Laplace priors. IEEE Trans. Image Process. **19**(1), 53–63 (2010)
10. DP Wipf, BD Rao, An empirical Bayesian strategy for solving the simultaneous sparse approximation problem. IEEE Trans. Signal Process **55**(7), 3704–3716 (2007)
11. SD Babacan, R Molina, AK Katsaggelos, Bayesian compressive sensing using Laplace priors. IEEE Trans. Image Process. **19**(1), 53–63 (2010)
12. JF Gu, SC Chan, WP Zhu, et al., Joint DOA estimation and source signal tracking with Kalman filtering and regularized QRD RLS algorithm. IEEE Trans Circuits. Syst. II, Exp Briefs **60**(1), 46–50 (2013)
13. B Yang, Projection approximation subspace tracking. IEEE Trans. Signal Process. **43**(1), 95–107 (Jan. 1995)
14. C Liu, G Wang, J Xin, et al., Low complexity subspace-based two-dimensional direction-of-arrivals tracking of multiple targets. Proc. ICSP, 1825–1829 (2012)
15. J Wu, J Xin, G Wang, et al., Two-dimensional direction tracking of coherent signals with two parallel uniform linear arrays. Proc. ICSP, 183–187 (2012)
16. N Vaswani, J Zhan, Recursive recovery of sparse signal sequences from compressive measurements: a review. IEEE Trans. Signal Process. **64**(13), 3523–3549 (2016)
17. N Vaswani, LS-CS-residual (LS-CS): compressive sensing on least squares residual. IEEE Trans. Signal Process. **58**(8), 4108–4120 (2010)
18. N Vaswani, W Lu, Modified-CS: modifying compressive sensing for problems with partially known support. IEEE Trans. Signal Process. **58**(9), 4595–4607 (2010)
19. M Friedlander, H Mansour, et al., Recovering compressively sampled signals using partial support information. IEEE Trans. Inf. Theory **58**(2), 1122–1134 (2012)
20. W Lu, N Vaswani, Regularized modified BPDN for noisy sparse reconstruction with partial erroneous support and signal value knowledge. IEEE Trans. Signal Process. **60**(1), 182–196 (2012)
21. XZ Gao et al., A sequential Bayesian algorithm for DOA tracking in time-varying environments. Chin. J. Electron. **24**(1), 140–145 (2015)
22. J Jia, L Yu, H Sun, et al., Dynamic DOA estimation via locally competitive algorithm. Proc. ICSP, 336–340 (2014)
23. Higuchi, Takuya, et al. Underdetermined blind separation and tracking of moving sources based on DOA-HMM. IEEE Int. Conf. Acoust, Speech and Signal Process, 3191-3195 (2014).
24. Y Tian, Z Chen, F Yin, Distributed imm-unscented kalman filter for speaker tracking in microphone array networks. IEEE Trans. Audio Speech Language Process. **23**(10), 1637–1647 (2015)
25. P Khomchuk, I Bilik, Dynamic direction-of-arrival estimation via spatial compressive sensing. Proc. Radar, 1191–1196 (2010)
26. M Hawes, L Mihaylova, F Septier, et al., A Bayesian compressed sensing Kalman filter for direction of arrival estimation. Proc. IEEE Int. Conf. Inf. Fusion, 969–975 (2015)
27. F Ning, L Ning, et al., Combining compressive sensing with particle filter for tracking moving wideband sound sources. IEEE Int. Conf. Signal Process, Commun and Comput, 1–6 (2015)
28. B Rafaely, Analysis and design of spherical microphone arrays. IEEE Trans. Speech, Audio Process. **13**(1), 135–143 (2005)
29. Q Huang, T Wang, Acoustic source localization in mixed field using spherical microphone arrays. EURASIP J. Adv. Sign. Process. **2014**(1), 1–16 (2014)
30. R Goossens, H Rogier, 2-D angle estimation with spherical arrays for scalar fields. Signal Process., IET **3**(3), 221–231 (2009)
31. H Sun, S Yan, U Svensson, Robust spherical microphone array beamforming with multiple-multi-null steering, and sidelobe control. in Proc. IEEE Workshop Appl. Signal Process. Audio Acoust, 113–116 (2009)
32. B. Rafaely, Fundamentals of Spherical Array Processing, Springer, chapter 1, pp. 10, 2015.
33. DA Linebarger, RD DeGroat, EM Dowling, Efficient direction-finding methods employing forward-backward averaging. IEEE Trans. Sign. Proc. **42**, 2136–2145 (1994)
34. A Roda, C Micheloni, Tracking sound sources by means of HMM. IEEE Int. Conf. Adv. Video and Signal-Based Surveillance, 83-85 (2011)
35. S Farahmand, GB Giannakis, G Leus, et al., Tracking target signal strengths on a grid using sparsity. Eurasip J. Adv. Sign. Proc. **2014**(1), 1–17 (2011)
36. C Bishop, *Pattern recognition and machine learning* (Springer-Verlag, New York, 2008), pp. 477–482
37. MJ Wainwright, MI Jordan, Graphical models, exponential families, and variational inference. Now Publishers Inc, pp. 159–164, 2008
38. EP Xing, MI Jordan, S Russell, A generalized mean field algorithm for variational inference in exponential families. Proceed. Conf. Uncertainty Artific. Intel, 583–591 (2012)
39. DG Tzikas, AC Likas, NP Galatsanos, The variational approximation for Bayesian inference. IEEE Sign. Proc. Magazine. **25**(6), 131–146 (2008)
40. P. Sebah, X. Gourdon, "Introduction to the Gamma Function," [Online]. Available: http://www.csie.ntu.edu.tw/~b89089/link/gammaFunction.pdf.
41. Z Yang, L Xie, C Zhang, Off-grid direction of arrival estimation using sparse Bayesian inference. IEEE Trans. Signal Process. **61**(1), 38–43 (2013)