

Off-grid fast relevance vector machine algorithm for direction of arrival estimation

ISSN 1751-8784

Received on 16th January 2015

Revised on 6th September 2015


Accepted on 29th September 2015

doi: 10.1049/iet-rsn.2015.0304

www.ietdl.org

Jincheng Lin , Xiaochuan Ma, Shefeng Yan, Chengpeng Hao, Geping Lin

Key Laboratory of Information Technology for Autonomous Underwater Vehicles, Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, People's Republic of China

 E-mail: ljcmym@163.com

Abstract: Direction of arrival (DOA) estimation is a basic and important problem in signal processing and has been widely applied. Its research has been advanced by the recently developed methods based on Bayesian compressive sensing (BCS). Among these methods, the ones combined with an off-grid (OG) model have been proved to be more accurate than the on-grid ones. However, the conventional BCS-based methods have a disadvantage of the slow speed. In this study, a high-efficiency iterative algorithm, based on the fast relevance vector machine and the OG model, is developed. This new approach applies to both the single- and multiple-snapshot cases. Numerical simulations show that the proposed method estimates DOAs more accurately than the ℓ_1 -penalisation method and computes more efficiently than the conventional BCS-based methods. Finally, comparisons with state-of-the-art methods and Cramer-Rao bound are also reported.

1 Introduction

Direction of arrival (DOA) estimation using sensor arrays is important in radar, sonar, seismic systems, acoustic source localisation, mobile communication etc. [1, 2]. The narrowband far-field source case is assumed in this paper, and the direction information is to be estimated. There are a lot of high-resolution DOA estimation algorithms such as MUSIC [3, 4] and ESPRIT [5]. However, most of these methods need a covariance matrix estimate, and rely on different prerequisites, for example, high signal-to-noise ratio (SNR) level, large number of snapshots and estimate of source number. In many practical applications, only a very small number of snapshots, even a single snapshot, are available for DOA estimation. Therefore, it is necessary to design estimators for the scenario of few snapshots (even one). Fortunately, a number of effective methods have been proposed for DOA estimation with low number of snapshots, for example, iterative adaptive approach (IAA)-based algorithms [6–10] and compressive sensing (CS)-based algorithms [10–12].

In [6], an IAA for amplitude and phase estimation (IAA-APES) has been proposed. The IAA-APES algorithm is an iterative and non-parametric algorithm that provides an accurate and high-efficiency estimate under severe snapshot limitations. The Bayesian information criterion [13] is used in conjunction with IAA-APES to give sparse results, which are usually assumed in DOA estimation. Furthermore, a parametric relaxation-based cyclic approach (RELAX) [14, 15] has also been used in IAA-APES&RELAX [6] to further improve the estimation accuracy. Actually, IAA-APES&RELAX is a high-efficiency off-grid (OG) spectral estimation algorithm. The comparisons between different OG algorithms will be demonstrated in our numerical simulations.

Recently, DOA estimation techniques have been advanced by the CS [16] methods. In the case of a single snapshot for DOA estimation, ℓ_1 -penalisation is a favourable approach to the sparse signal recovery, because it does not depend on the sample covariance matrix and sources' correlativity, and can accurately estimate multiple DOAs only in a single snapshot [11].

In the CS field, sparse Bayesian learning (SBL)/inference (SBI) with the relevance vector machine (RVM) [17] has been another popular method for the sparse signal recovery. The concept of Bayesian CS (BCS) has been proposed in [18]. In BCS, the CS

inversion problem is formulated from an SBL perspective. The analysis of RVM [19, 20] has proved that the RVM provides a tighter approximation to the ℓ_0 -norm sparsity than the ℓ_1 -norm.

The approach using BCS in DOA estimation has been presented in [21], including single-snapshot BCS [18] and multiple-snapshot BCS (MT-BCS) [22]. Some RVM-based DOA estimation methods are also researched in [23–30]. In the case of multiple snapshots, the SBL approach for the ℓ_1 -SVD model [11] has been proposed in [31]. These methods all show the accurate and sparse results of the DOA estimation problem. Nevertheless, only the conventional RVM has been used in these articles, which is involved in inverting a large matrix and consumes a large amount of computation. Compared with the conventional RVM algorithm, the fast RVM algorithm developed in [32, 33] can compute more efficiently. Through adding, deleting and re-estimation candidate basis functions in each iteration, the fast RVM can monotonically maximise the marginal likelihood and choose basis functions smarter.

Although the existing CS or BCS-based methods have shown their outstanding performance in DOA estimation, there are still more or less deviations when the actual DOAs are not on the sampling grid. Not to be constrained on the sampling grid, OG methods for DOA estimation are proposed in [34]. There are also some further discussion about the OG method in [35–38]. In [31], an SBL method based on the OG model for DOA estimation has been proposed. This method is referred as OG-SBI. It can obtain a more accurate estimation not restricted by the fixed sampling grid (on-grid model) and an excellent sparsity as BCS.

Nevertheless, the OG-SBI algorithm is realised by the conventional RVM, there is no advantage in computational efficiency. In this paper, we propose a fast RVM algorithm using the OG adjustment for both single- and MT DOA estimation. The proposed algorithm has the advantages of both the accurate estimation by OG model and the high-efficiency computation by fast RVM. We refer to the proposed algorithm in this paper as OG fast RVM (OG-FastRVM). Some state-of-the-art single-snapshot algorithms and subspace-based algorithms are taken as comparisons in numerical simulations. Through numerical simulations, we show that OG-FastRVM has a smaller root mean square error (RMSE) in comparison with the ℓ_1 -penalisation on a fixed grid. The simulations also demonstrate the advantages of OG-FastRVM compared with the state-of-the-art OG algorithms,

for example, OG-SBI and IAA-APES&RELAX. In addition, the estimation accuracy of the MT version of OG-FastRVM (MT-OG-FastRVM) improves with the increase of the number of snapshots.

The rest of this paper is organised as follows. Section 2 introduces the sparse representation for the DOA estimation model. Section 3 studies the SBI of the real-valued model for DOA estimation and the origination of fast RVM algorithm. Section 4 introduces the proposed OG-FastRVM algorithm and its MT version. Section 5 presents our numerical simulations. Section 6 concludes this paper.

2 Sparse representation for the DOA estimation model

In this paper, we consider K far-field narrowband signals $s_k(t)$, $k = 1, \dots, K$, impinging on a uniform linear array (ULA) of M omnidirectional sensors. The signal directions are θ_k , $k = 1, \dots, K$. The observation model can be presented as

$$\mathbf{y}(t) = \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t) + \mathbf{n}(t), \quad (1)$$

where $\mathbf{y}(t) = [y_1(t), \dots, y_M(t)]^T$, $\mathbf{s}(t) = [s_1(t), \dots, s_K(t)]^T$, $\boldsymbol{\theta} = [\theta_1, \dots, \theta_K]^T$ and $\mathbf{n}(t) = [n_1(t), \dots, n_M(t)]^T$. Precisely, $y_m(t)$ and $n_m(t)$, $m = 1, \dots, M$, are the output and noise of m th sensor at time t , respectively. The matrix $\mathbf{A}(\boldsymbol{\theta})$ is the so-called array manifold matrix, which is composed of the K steering vectors $\mathbf{a}(\theta_k)$, $k = 1, \dots, K$.

Here, we formulate the DOA estimation problem as a sparse representation problem. The single-snapshot formulation in this section parallels the ones in [11, 12], where it was presented as the application of the ℓ_1 -norm penalisation algorithm. The single-snapshot case is first considered and to be extended to the multi-snapshot one in Section 4.2. To handle this problem as a sparse representation problem, we introduce an overcomplete representation $\tilde{\mathbf{A}}$ in terms of N DOAs of interest. Let $\tilde{\boldsymbol{\theta}} = \{\tilde{\theta}_1, \dots, \tilde{\theta}_N\}$ be a fixed sampling grid which covers all the DOAs of the K actual signals. The number of potential DOAs N denotes the grid number and typically satisfies $N \gg M > K$. Without loss of generality, let $\tilde{\boldsymbol{\theta}}$ be a uniform grid, the grid interval $r = \tilde{\theta}_n - \tilde{\theta}_{n-1} = \pi/N$.

We denote the matrix \mathbf{A} composed of N potential steering vectors: $\mathbf{A} = [\mathbf{a}(\tilde{\theta}_1), \mathbf{a}(\tilde{\theta}_2), \dots, \mathbf{a}(\tilde{\theta}_N)]$.

The observed single-snapshot signal is expressed as

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}, \quad (2)$$

where \mathbf{x} is the signal vector. For $n = 1, \dots, N$, \mathbf{x} has K non-zero elements

$$\begin{aligned} x_{n_k} &= s_k(t), & k = 1, \dots, K; \\ x_n &= 0, & \text{otherwise,} \end{aligned} \quad (3)$$

with $n_k \in \{1, \dots, N\}$. Assuming $\tilde{\theta}_{n_k}$, $k = 1, \dots, K$ belongs to the set of the K DOAs θ_k , $k = 1, \dots, K$, and we will discuss the biased assumption in Section 4 in which an OG method will be used to offset the bias. Therefore, the problem turns to be CS inversion problem for recovering the vector \mathbf{x} due to the sparsity of $N \gg K$. A typical means of solving such a problem is via an ℓ_1 -regularised formulation [11, 39]

$$\arg \min_{\mathbf{x}} \{ \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1 \}, \quad (4)$$

where the parameter λ controls the relative weight applied to the Euclidian error and the signal sparsity. Instead, the constrained version of the aforementioned problem is shown as follows:

$$\begin{aligned} \min & \quad \|\mathbf{x}\|_1 \\ \text{s.t.} & \quad \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 \leq \epsilon, \end{aligned} \quad (5)$$

where ϵ is a user specified parameter, it should be chosen high enough so that the probability of $\|\mathbf{n}\|_2^2 \geq \epsilon$ is small. The problem in (5) can be efficiently solved in the framework of second-order cone [11].

3 Sparse Bayesian inference

To deal with the complex data model (4) via the BCS approach [21], (2) is rewritten as

$$\begin{bmatrix} \Re\{\mathbf{y}\} \\ \Im\{\mathbf{y}\} \end{bmatrix} = \begin{bmatrix} \Re\{\mathbf{A}\} & -\Im\{\mathbf{A}\} \\ \Im\{\mathbf{A}\} & \Re\{\mathbf{A}\} \end{bmatrix} \begin{bmatrix} \Re\{\mathbf{x}\} \\ \Im\{\mathbf{x}\} \end{bmatrix} + \begin{bmatrix} \Re\{\mathbf{n}\} \\ \Im\{\mathbf{n}\} \end{bmatrix}, \quad (6)$$

where $\Re\{\cdot\}$ and $\Im\{\cdot\}$ denote the real and imaginary parts, respectively. Therefore, the complex problem is turned into a real-valued form in order that we can deal with it through the BCS method and further the fast RVM algorithm. The compact form of (6) is shown as follows:

$$\tilde{\mathbf{y}} = \tilde{\mathbf{A}}\tilde{\mathbf{x}} + \tilde{\mathbf{n}}. \quad (7)$$

Under an assumption of circular-symmetric Gaussian noise [40], we have the Gaussian likelihood model of (7)

$$p(\tilde{\mathbf{y}}|\tilde{\mathbf{x}}, \sigma^2) = (2\pi\sigma^2)^{-N/2} \exp\left\{-\frac{1}{2\sigma^2}|\tilde{\mathbf{y}} - \tilde{\mathbf{A}}\tilde{\mathbf{x}}|^2\right\}. \quad (8)$$

The aforementioned discussion converted the conventional CS inversion problem into a maximum *a posteriori* approximation to a Bayesian linear regression, while the sparsity of $\tilde{\mathbf{x}}$ is limited by a Laplace sparseness prior. Since the Laplace prior is not conjugate to the Gaussian likelihood [18], a more flexible prior model needs to be invoked. This issue has been referred in RVM [17]. In the RVM, a hierarchical prior on $\tilde{\mathbf{x}}$ has been proposed, which allows convenient conjugate-exponential analysis. The subsequent theoretical analysis for RVM [19, 20] has proved that the RVM provides a tighter approximation to the ℓ_0 -norm sparsity than the ℓ_1 -norm. The two-stage hierarchical prior on $\tilde{\mathbf{x}}$ is

$$\begin{aligned} p(\tilde{\mathbf{x}}|\boldsymbol{\alpha}) &= \prod_{i=1}^N \mathcal{N}(w_i|0, \alpha_i^{-1}), \\ p(\boldsymbol{\alpha}|a, b) &= \prod_{i=1}^N \Gamma(\alpha_i|a, b), \end{aligned} \quad (9)$$

where $\boldsymbol{\alpha}$ is the vector of N hyperparameters, each α_i indicates the precision of a Gaussian density associated with \tilde{x}_i .

On the basis of the combination of (8) and (9), given $\tilde{\mathbf{y}}$, our objective is to evaluate $\tilde{\mathbf{x}}$ and $\sigma^{-2}(\alpha_0)$: namely, maximising the posterior density function

$$p(\tilde{\mathbf{x}}, \boldsymbol{\alpha}, \alpha_0|\tilde{\mathbf{y}}) = p(\tilde{\mathbf{x}}|\tilde{\mathbf{y}}, \boldsymbol{\alpha}, \alpha_0)p(\boldsymbol{\alpha}, \alpha_0|\tilde{\mathbf{y}}). \quad (10)$$

According to Bayes' rule, (10) turns to be

$$\max\{[p(\tilde{\mathbf{x}}|\tilde{\mathbf{y}}, \boldsymbol{\alpha}, \alpha_0)] \cdot [p(\tilde{\mathbf{y}}|\boldsymbol{\alpha}, \alpha_0)p(\boldsymbol{\alpha})p(\alpha_0)]\}. \quad (11)$$

Given the hyperparameters $\boldsymbol{\alpha}$ and α_0 , by maximising the first term of (11), the mean and covariance of $\tilde{\mathbf{x}}$ from the posterior expression are evident [17, 18]

$$\begin{aligned} \tilde{\boldsymbol{\mu}} &= \alpha_0 \tilde{\mathbf{A}}^T \tilde{\mathbf{y}}, \\ \tilde{\boldsymbol{\Sigma}} &= (\alpha_0 \tilde{\mathbf{A}}^T \tilde{\mathbf{A}} + \mathbf{\Lambda})^{-1}, \end{aligned} \quad (12)$$

where $\mathbf{\Lambda} = \text{diag}(\alpha_1, \alpha_2, \dots, \alpha_N)$. For the second term of (11), fixing $\tilde{\boldsymbol{\mu}}$ and $\tilde{\boldsymbol{\Sigma}}$, the logarithm marginal likelihood for $\boldsymbol{\alpha}$ and α_0 , $\mathcal{L}(\boldsymbol{\alpha}, \alpha_0)$

can be expressed as follows [18]:

$$\mathcal{L}(\boldsymbol{\alpha}, \alpha_0) = -\frac{1}{2}[M \log 2\pi + \log |\mathbf{C}| + \tilde{\mathbf{y}}^T \mathbf{C}^{-1} \tilde{\mathbf{y}}], \quad (13)$$

with $\mathbf{C} = \alpha_0^{-1} \mathbf{I} + \tilde{\mathbf{A}} \boldsymbol{\Lambda}^{-1} \tilde{\mathbf{A}}^T$. The conventional method employs the estimates for $\boldsymbol{\alpha}$ and α_0 to solve the following partial differential equations [17]

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \alpha_i} &\Rightarrow \alpha_i^{\text{new}} = \frac{\gamma_i}{\mu_i^2}, \quad i \in \{1, 2, \dots, N\} \\ \frac{\partial \mathcal{L}}{\partial \alpha_0} &\Rightarrow \frac{1}{\alpha_0^{\text{new}}} = \sigma^2 = \frac{|\tilde{\mathbf{y}} - \tilde{\mathbf{A}} \tilde{\boldsymbol{\mu}}|^2}{M - \sum_i \gamma_i} \end{aligned} \quad (14)$$

where $\gamma_i = 1 - \alpha_i \boldsymbol{\Sigma}_{ii}$, with $\boldsymbol{\Sigma}_{ii}$ the i th diagonal element of $\boldsymbol{\Sigma}$ in (12).

Noting that the combination of (12) and (14) implies an iterative algorithm, that is, calculating $\{\tilde{\boldsymbol{\mu}}, \boldsymbol{\Sigma}\}$ for the given $\{\boldsymbol{\alpha}, \alpha_0\}$, vice versa. When the iteration reaches a convergence criterion, we can get the estimation of sparse signals as

$$\hat{\mathbf{x}} = \boldsymbol{\mu}, \quad (15)$$

whose K largest terms are $\{\hat{x}_{i_1}, \dots, \hat{x}_{i_K}\}$. Here, $\tilde{\boldsymbol{\mu}} = [\Re\{\boldsymbol{\mu}\}^T \Im\{\boldsymbol{\mu}\}^T]^T$. Therefore, the corresponding K underlying DOAs are $\{\tilde{\theta}_{i_1}, \dots, \tilde{\theta}_{i_K}\}$.

However, solving the matrix inversion in (12) consumes a large amount of computation, especially in the case of a large dimension N : namely, a dense sampling grid [18]. The numerical experiment in [31] has shown this drawback. For computing more efficiently, the fast RVM algorithm has been developed in [32, 33], and applied in BCS [18] and pattern synthesis [41]. The original idea of the fast RVM algorithm is extracting the part affected by α_i from $\mathcal{L}(\boldsymbol{\alpha}; \alpha_0)$, given α_0

$$\mathcal{L}(\boldsymbol{\alpha}; \alpha_0) = \mathcal{L}(\boldsymbol{\alpha}_{-i}; \alpha_0) + \ell(\alpha_i; \alpha_0), \quad (16)$$

where $\mathcal{L}(\boldsymbol{\alpha}_{-i}; \alpha_0)$ is the marginal likelihood with α_i (and thus $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{a}}_i$) removed from the model. Therefore, the terms in α_i can be independently dealt with, thereby we can sequentially add and delete the candidate basis function (one column of $\tilde{\mathbf{A}}$) or re-estimate the value of α_i [32, 33]. The criterion of which the adding, deleting or re-estimating branch should be chosen is maximising the increment $\Delta \mathcal{L}$ of the marginal likelihood $\mathcal{L}(\boldsymbol{\alpha}; \alpha_0)$. The details are appended in Algorithms 3, 4 and 6 (Figs. 1a, 12 and 14).

The fast RVM algorithm enables an efficiently sequential addition and deletion of candidate basis function (columns of $\tilde{\mathbf{A}}$) to monotonically maximise the marginal likelihood $\mathcal{L}(\boldsymbol{\alpha}; \alpha_0)$. Compared with the original iteration presented in (12) and (14), the fast RVM sequentially adds (or deletes) candidate basis function to the procedure until all K relevant sparse basis functions have been included. The weights associated with whose basis functions are non-zero. Therefore, the computational load of the algorithm is more related to K than N . Especially, when the underlying signal is indeed sparse ($K \ll N$), the fast RVM is more efficient than the original version.

In Algorithms 1–6 (Figs. 1, 10–14), we will show the implementation details of the fast RVM algorithm with the OG adjustment technique embedded in it. Algorithm 1 (Fig. 1) in Section 4 is the overall structure of the proposed algorithm and Algorithm 2–6 (Figs. 10–14) in the Appendix are the details.

4 OG-FastRVM algorithm

4.1 Theoretical formulation

An assumption has been considered in the aforementioned formulation that the underlying DOAs θ_k , $k=1, \dots, K$ should

Data: $\tilde{\boldsymbol{\Phi}}, \tilde{\mathbf{y}}$

Result: $\tilde{\boldsymbol{\mu}}, \boldsymbol{\Sigma}, \sigma^2$

```

1 initialization  $\tilde{\boldsymbol{\mu}}, \boldsymbol{\Sigma}, \sigma^2, \tilde{\boldsymbol{\Phi}}, \tilde{\alpha}_i, s_m, q_m$ ; /*  $s_m$  and
 $q_m$  are defined in Algorithm 2 */
2 repeat
3   select a  $\tilde{\boldsymbol{\phi}}_i$  for  $i = 1 : 2N$ ;
4   compute  $\theta_i \triangleq q_i^2 - s_i$ ;
5   if  $\theta_i > 0$  then
6     if  $\tilde{\alpha}_i < \infty$  then
7       re-estimate  $\tilde{\alpha}_i$ ; /* Algorithm 4 */
8       calculate  $\beta_i$ ; /* Algorithm 5 */
9     else
10      update  $\tilde{\alpha}_i$  and add  $\tilde{\boldsymbol{\phi}}_i$ ; /* Algorithm
11      3 */
12   else
13     if  $\tilde{\alpha}_i < \infty$  then
14       delete  $\tilde{\boldsymbol{\phi}}_i$  and set  $\tilde{\alpha}_i = \infty$ ;
15       /* Algorithm 6 */
16   update the noise level  $\sigma^2$ ;
17   update  $\tilde{\boldsymbol{\mu}}, \boldsymbol{\Sigma}, s_m, q_m$ ;
18 until converged;
```

Fig. 1 OG-FastRVM

locate just on the sampling grids $\tilde{\boldsymbol{\theta}} = \{\tilde{\theta}_1, \dots, \tilde{\theta}_N\}$. However, there is always more or less bias between the actual θ_k and the nearest sampling grid point $\tilde{\theta}_{n_k}$.

Considering an actual DOA $\theta_k \notin \{\tilde{\theta}_1, \dots, \tilde{\theta}_N\}$, the nearest DOA sampling is $\tilde{\theta}_{n_k}$, $n_k \in \{1, \dots, N\}$. Compared with the direct expression $\mathbf{a}(\theta_k) \simeq \mathbf{a}(\tilde{\theta}_{n_k})$, the more accurate linear approximation can be formulated

$$\mathbf{a}(\theta_k) \simeq \mathbf{a}(\tilde{\theta}_{n_k}) + \mathbf{b}(\tilde{\theta}_{n_k})(\theta_k - \tilde{\theta}_{n_k}), \quad (17)$$

with $\mathbf{b}(\tilde{\theta}_{n_k}) = d\mathbf{a}(\tilde{\theta}_{n_k})/d\tilde{\theta}_{n_k} = \mathbf{a}'(\tilde{\theta}_{n_k})$. On the basis of this consideration, the single-snapshot observation model in (2) is re-presented as [31, 34]

$$\begin{aligned} \mathbf{y} &= [\mathbf{A} + \mathbf{B} \text{diag}(\boldsymbol{\beta})] \mathbf{x} + \mathbf{n} \\ &= \boldsymbol{\Phi}(\boldsymbol{\beta}) \mathbf{x} + \mathbf{n}, \end{aligned} \quad (18)$$

where $\mathbf{B} = [\mathbf{b}(\tilde{\theta}_1), \dots, \mathbf{b}(\tilde{\theta}_N)]$, $\boldsymbol{\beta} = [\beta_1, \dots, \beta_N]^T$ and $\boldsymbol{\Phi}(\boldsymbol{\beta}) = [\boldsymbol{\phi}_1, \dots, \boldsymbol{\varphi}_N]$. If there is a target direction θ_k located near $\tilde{\theta}_{n_k}$, the corresponding β_{n_k} should get a value $\in [-(1/2)r, (1/2)r]$, with r as the grid interval. Otherwise, $\beta_n = 0$. To get a more accurate estimation without depending on the sampling grid density, $\boldsymbol{\beta}$ needs to be estimated.

Following the definition in (6), the real-valued model of (18) is:

$$\tilde{\mathbf{y}} = \tilde{\boldsymbol{\Phi}}(\boldsymbol{\beta}) \tilde{\mathbf{x}} + \tilde{\mathbf{n}}. \quad (19)$$

Then the processed dimensions M and N turn to be $2M$ and $2N$. The observation model used in the fast RVM algorithm is real-valued. When we deal with a basis function $\tilde{\boldsymbol{\phi}}_i$, if the adjustment factor is estimated, the conjugate basis function $\tilde{\boldsymbol{\phi}}_{i+N}$ (or $\tilde{\boldsymbol{\phi}}_{i-N}$) will be affected. Therefore, the OG method used in [31] is not available here. In this section, a modified OG method aimed to the fast RVM algorithm will be proposed.

When $\boldsymbol{\beta} = [\beta_1, \dots, \beta_N]^T$ is estimated, it is evident that we need to estimate only β_{n_k} , $k \in \{1, \dots, K\}$ associated with the K sampling grids $\tilde{\theta}_{n_k}$ which are closest to the signal DOAs $\{\theta_1, \dots, \theta_K\}$. In practical calculation, the basis function set $\{\boldsymbol{\phi}_i\}$, where we choose

the estimation index i , should satisfy

$$\phi_i \in \{\phi_m | \arg_m \{\mu_m \in \text{the maximum } K \text{ entries of } \boldsymbol{\mu}\}\},$$

for the real-valued form

$$\tilde{\phi}_i \in \{\tilde{\phi}_m | \arg_m \{\tilde{\mu}_m \in \text{the maximum } 2K \text{ entries of } \tilde{\boldsymbol{\mu}}\}\}. \quad (20)$$

In one iteration, if the algorithm [detailed in Algorithm 1 (Fig. 1)] runs into the re-estimating branch for $\tilde{\mu}_i$ (thus $\tilde{\alpha}_i$) and the condition (20) is satisfied, we will estimate β_i . We estimate $\boldsymbol{\beta}$ only in the re-estimate branch, and only a single $\beta_i = \beta_i + j0$ is estimated (due to the representation for only the deviation, β_i is real-valued). Therefore, the change of $\boldsymbol{\Phi}$ caused by $\boldsymbol{\beta}$ is associated with only the column ϕ_i . Similar to that, the change of $\tilde{\boldsymbol{\Phi}}$ happens at the i th column $\begin{bmatrix} \Delta\phi_i^R \\ \Delta\phi_i^I \end{bmatrix}$ and the $(N+i)$ th column $\begin{bmatrix} -\Delta\phi_i^I \\ \Delta\phi_i^R \end{bmatrix}$. Here, $(\cdot)^R$ and $(\cdot)^I$ denote the real and imaginary parts, respectively.

According to the index i , the relationship between the corresponding vectors is

$$\phi_i = \mathbf{a}_i + \mathbf{b}_i \beta_i,$$

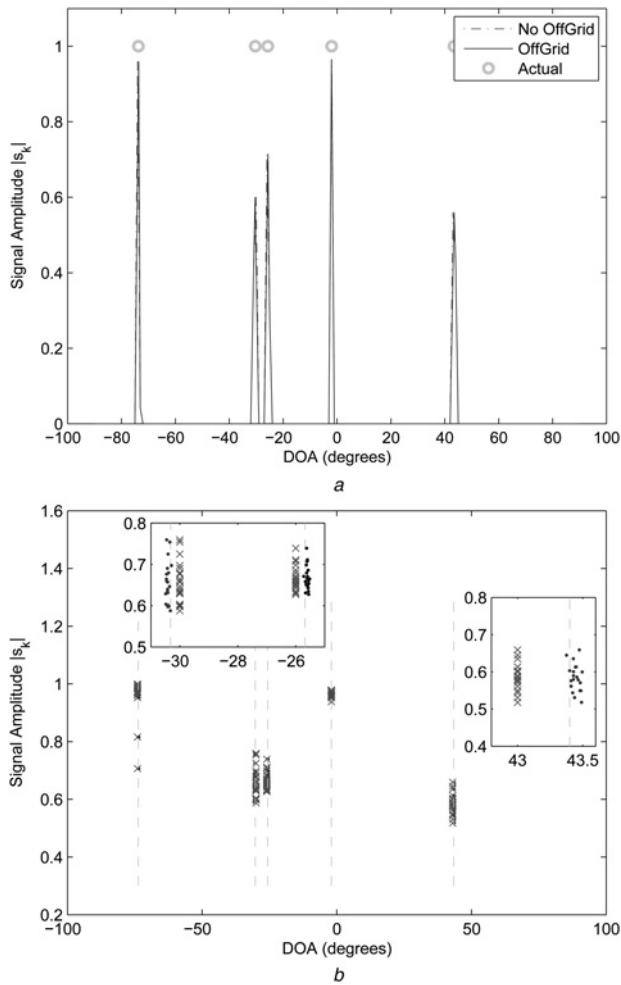


Fig. 2 Single sample DOA estimation using FastRVM and OG-FastRVM as# Overall

b Detail illustrations: SNR = 20dB, $K = 5$, $\Delta\theta_{\min} = 4.64^\circ$, 20 trials
In (b), dot and cross represent OG-FastRVM and FastRVM, respectively, and the dashed-line represents the actual position of DOAs

with the real-valued form

$$\begin{aligned} \begin{bmatrix} \phi_i^R & -\phi_i^I \\ \phi_i^I & \phi_i^R \end{bmatrix} &= \begin{bmatrix} \mathbf{a}_i^R & -\mathbf{a}_i^I \\ \mathbf{a}_i^I & \mathbf{a}_i^R \end{bmatrix} + \begin{bmatrix} \Re(\mathbf{b}_i \beta_i) & -\Im(\mathbf{b}_i \beta_i) \\ \Im(\mathbf{b}_i \beta_i) & \Re(\mathbf{b}_i \beta_i) \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{a}_i^R & -\mathbf{a}_i^I \\ \mathbf{a}_i^I & \mathbf{a}_i^R \end{bmatrix} + \begin{bmatrix} \mathbf{b}_i^R & -\mathbf{b}_i^I \\ \mathbf{b}_i^I & \mathbf{b}_i^R \end{bmatrix} \begin{bmatrix} \beta_i & 0 \\ 0 & \beta_i \end{bmatrix}. \end{aligned} \quad (21)$$

Our target is to calculate the $\boldsymbol{\beta}$ in order to minimise the deviation between the observation and the model in (18). The target function is

$$\arg \min_{\boldsymbol{\beta}} \|\mathbf{y} - [\mathbf{A} + \mathbf{B} \text{diag}(\boldsymbol{\beta})] \mathbf{x}\|_2^2. \quad (22)$$

Here, we simplify it as follows:

$$\begin{aligned} &\|\mathbf{y} - [\mathbf{A} + \mathbf{B} \text{diag}(\boldsymbol{\beta})] \mathbf{x}\|_2^2 \\ &\simeq \|\mathbf{y} - [\mathbf{A} + \mathbf{B} \text{diag}(\boldsymbol{\beta})] \boldsymbol{\mu}\|_2^2 \\ &= \|\mathbf{y} - \mathbf{A} \boldsymbol{\mu} - \beta_i \mathbf{b}_i \boldsymbol{\mu}_i\|_2^2 \\ &= \left\| \tilde{\mathbf{y}} - \tilde{\mathbf{A}} \tilde{\boldsymbol{\mu}} - \begin{bmatrix} \mathbf{b}_i^R & -\mathbf{b}_i^I \\ \mathbf{b}_i^I & \mathbf{b}_i^R \end{bmatrix} \begin{bmatrix} \beta_i & 0 \\ 0 & \beta_i \end{bmatrix} \begin{bmatrix} \boldsymbol{\mu}_i^R \\ \boldsymbol{\mu}_i^I \end{bmatrix} \right\|_2^2 \\ &\simeq \|\tilde{\mathbf{y}} - \tilde{\mathbf{A}} \tilde{\boldsymbol{\mu}} - \beta_i \tilde{\mathbf{B}} \tilde{\boldsymbol{\mu}}_i\|_2^2 \\ &= \|\tilde{\mathbf{n}} - \beta_i \tilde{\mathbf{B}} \tilde{\boldsymbol{\mu}}_i\|_2^2, \end{aligned} \quad (23)$$

where $\tilde{\boldsymbol{\mu}}_i = [\tilde{\mu}_i, \tilde{\mu}_{i+N}]^T = [\mu_i^R, \mu_i^I]^T$ and $\tilde{\boldsymbol{\mu}} = [\tilde{\mu}_1, \dots, \tilde{\mu}_{2K}]$ $\{i_1, \dots, i_{2K}\}$ satisfy (20). Furthermore

$$\begin{aligned} \|\tilde{\mathbf{n}} - \beta_i \tilde{\mathbf{B}} \tilde{\boldsymbol{\mu}}_i\|_2^2 &= (\tilde{\mathbf{n}} - \beta_i \tilde{\mathbf{B}} \tilde{\boldsymbol{\mu}}_i)^T (\tilde{\mathbf{n}} - \beta_i \tilde{\mathbf{B}} \tilde{\boldsymbol{\mu}}_i) \\ &= \beta_i^2 \tilde{\boldsymbol{\mu}}_i^T \tilde{\mathbf{B}}^T \tilde{\mathbf{B}} \tilde{\boldsymbol{\mu}}_i - 2\beta_i \tilde{\mathbf{n}}^T \tilde{\mathbf{B}} \tilde{\boldsymbol{\mu}}_i + \tilde{\mathbf{n}}^T \tilde{\mathbf{n}}. \end{aligned} \quad (24)$$

Therefore, (22) is approximately equivalent to

$$\beta_i = \arg \min_{\beta_i \in \left[-\frac{1}{2}r, \frac{1}{2}r\right]} \{\beta_i^2 \tilde{\boldsymbol{\mu}}_i^T \tilde{\mathbf{B}}^T \tilde{\mathbf{B}} \tilde{\boldsymbol{\mu}}_i - 2\beta_i \tilde{\mathbf{n}}^T \tilde{\mathbf{B}} \tilde{\boldsymbol{\mu}}_i\}, \quad (25)$$

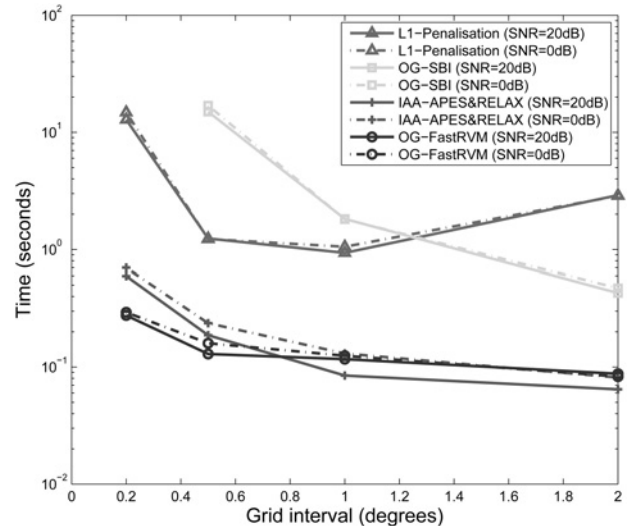


Fig. 3 Time consumptions of four methods: $M = 24$, $K = 5$

thus

$$\beta_i = \begin{cases} \frac{\bar{\boldsymbol{\mu}}^T \bar{\mathbf{B}} \boldsymbol{\mu}_i}{\boldsymbol{\mu}_i^T \bar{\mathbf{B}} \boldsymbol{\mu}_i} = \frac{(\tilde{\mathbf{y}} - \bar{\mathbf{A}} \bar{\boldsymbol{\mu}})^T \bar{\mathbf{B}} \boldsymbol{\mu}_i}{\boldsymbol{\mu}_i^T \bar{\mathbf{B}} \boldsymbol{\mu}_i}, & \beta_i \in \left[-\frac{1}{2}r, \frac{1}{2}r\right] \\ 0, & \text{otherwise} \end{cases} \quad (26)$$

The Algorithm 1 (Fig. 1) is the overall structure of the OG fast RVM (OG-FastRVM) algorithm. The details are shown in Appendix.

4.2 Extend to the MT model

Considering a MT model of T snapshots, the signal model of (7) turns into

$$\tilde{\mathbf{y}}_t = \tilde{\mathbf{A}} \tilde{\mathbf{x}}_t + \tilde{\mathbf{n}}_t, \quad t = 1, \dots, T. \quad (27)$$

Different from the single-snapshot BCS algorithm, MT-BCS [22] exploits the correlation of the noise strength of multiple snapshots. The MT version of the marginal likelihood (13), or equivalently, its logarithm is expressed as follows:

$$\begin{aligned} \mathcal{L}_{MT}(\boldsymbol{\alpha}, \alpha_0) &= \sum_{t=1}^T \log p(\tilde{\mathbf{y}}_t | \boldsymbol{\alpha}, \alpha_0) \\ &= \sum_{t=1}^T \log \int p(\tilde{\mathbf{y}}_t | \tilde{\mathbf{x}}_t, \alpha_0) p(\tilde{\mathbf{x}}_t | \boldsymbol{\alpha}) d\tilde{\mathbf{x}}_t \\ &= -\frac{1}{2} \sum_{t=1}^T [2M \log 2\pi + \log |\mathbf{C}_{MT}| + \tilde{\mathbf{y}}_t^T \mathbf{C}_{MT}^{-1} \tilde{\mathbf{y}}_t], \end{aligned} \quad (28)$$

where $\mathbf{C}_{MT} = \alpha_0^{-1} \mathbf{I} + \tilde{\mathbf{A}} \boldsymbol{\Lambda}^{-1} \tilde{\mathbf{A}}^T$. The following procedures resemble the OG-FastRVM algorithm. Thus, the estimated sparse signal

vector $\tilde{\boldsymbol{\mu}}_{MT}$ is calculated as follows

$$\tilde{\boldsymbol{\mu}}_{MT} = \frac{1}{T} \sum_{t=1}^T \left[\mathbf{0}_{(0)} \tilde{\mathbf{A}}^T \tilde{\mathbf{A}} + \boldsymbol{\Lambda} \right]^{-1} \tilde{\mathbf{A}}^T \tilde{\mathbf{y}}_t. \quad (29)$$

Considering the condition of multiple snapshots, the OG parameters $\boldsymbol{\beta}_t$ corresponding to different snapshots is determined by

$$\boldsymbol{\beta}_t = \arg \min_{\boldsymbol{\beta}_t} \|\mathbf{y}_t - [\mathbf{A} + \mathbf{B} \text{diag}(\boldsymbol{\beta}_t)] \boldsymbol{\mu}_{MT}\|_2^2. \quad (30)$$

Therefore, the K potential values of DOAs is

$$\hat{\boldsymbol{\theta}}_m = \frac{1}{T} \sum_{t=1}^T (\tilde{\boldsymbol{\theta}}_m + \boldsymbol{\beta}_{t,m}) \quad (31)$$

$$\text{s.t. } \boldsymbol{\mu}_m \in \{\text{the maximum } K \text{ entries of } \boldsymbol{\mu}_{MT}\}.$$

It should be noted that $\tilde{\boldsymbol{\mu}}_{MT} \triangleq [\Re\{\boldsymbol{\mu}_{MT}\}^T \Im\{\boldsymbol{\mu}_{MT}\}^T]^T$ and $\tilde{\mathbf{y}}_t \triangleq [\Re\{\mathbf{y}_t\}^T \Im\{\mathbf{y}_t\}^T]^T$.

5 Numerical simulations

5.1 Single-snapshot OG-FastRVM

In this section, a few numerical simulations are presented to show the performance of the proposed OG-FastRVM method. Compared with the state-of-the-art methods for single-snapshot DOA estimation, its advantages and drawbacks will be pointed out. An ULA with half-wavelength inter-element spacing is considered. The origin is set at the midpoint of the ULA. A uniform sampling grid $\{-90^\circ, -90^\circ + r, \dots, 90^\circ - r, 90^\circ\}$ is adopted with r the DOA sampling interval. In this section, we use only the single-snapshot model in the simulations. We take ℓ_1 -penalisation [11], OG-SBI [31] and IAA-APES&RELAX [6] as three representative methods for single snapshot, and as comparisons with our proposed method. ℓ_1 -penalisation is an on-grid method, whereas OG-SBI and IAA-APES&RELAX are two OG methods.

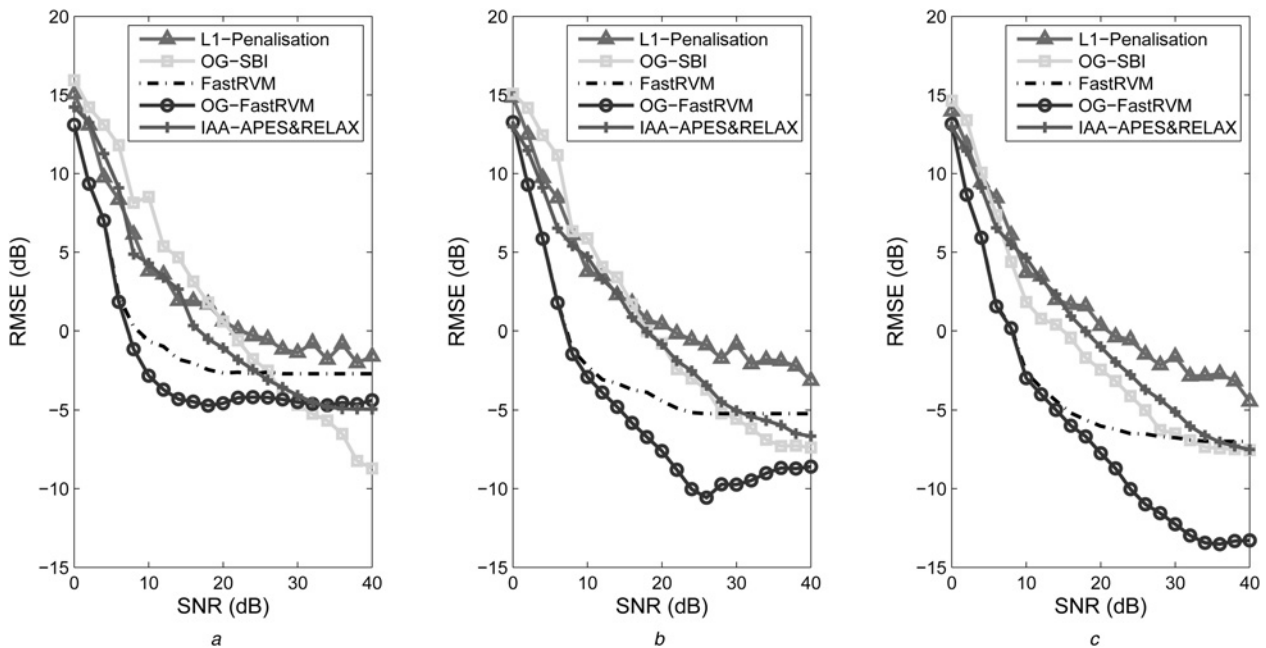


Fig. 4 RMSEs of ℓ_1 -penalisation, OG-SBI, FastRVM, OG-FastRVM and IAA-APES&RELAX versus SNR for various r

a $r = 2^\circ$
 b $r = 1^\circ$
 c $r = 0.5^\circ, K = 2, \Delta\theta = 15^\circ, M = 8, 200$ trials

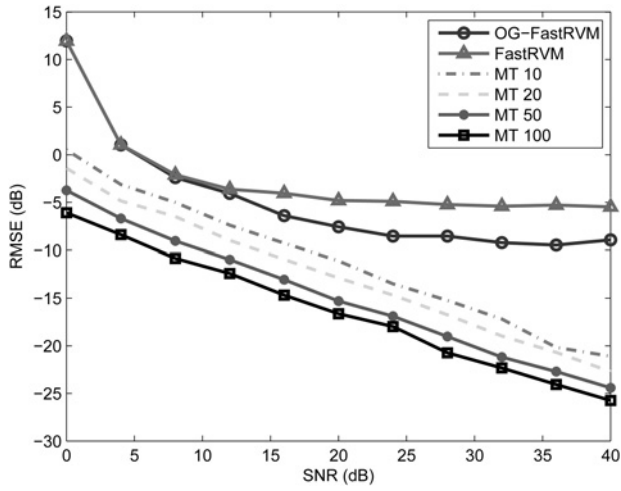


Fig. 5 RMSEs of OG-FastRVM and MT-OG-FastRVM against SNR: $K = 2$, $M = 8$, 200 trials

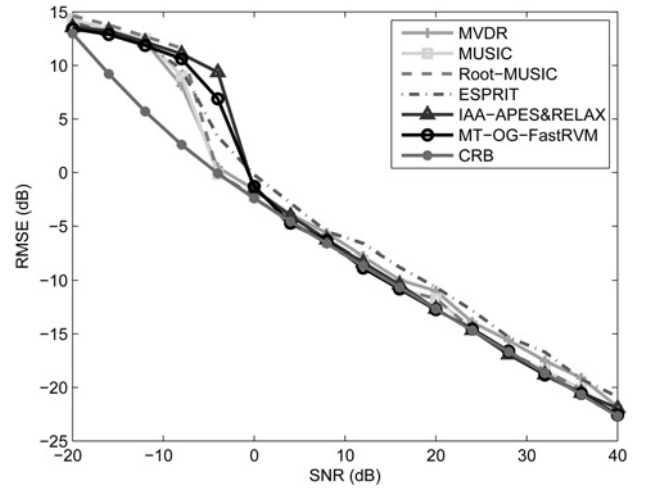


Fig. 8 RMSEs of different algorithms for DOA estimation and CRB against SNR: $T = 20$, $K = 2$, $M = 8$, 200 trials

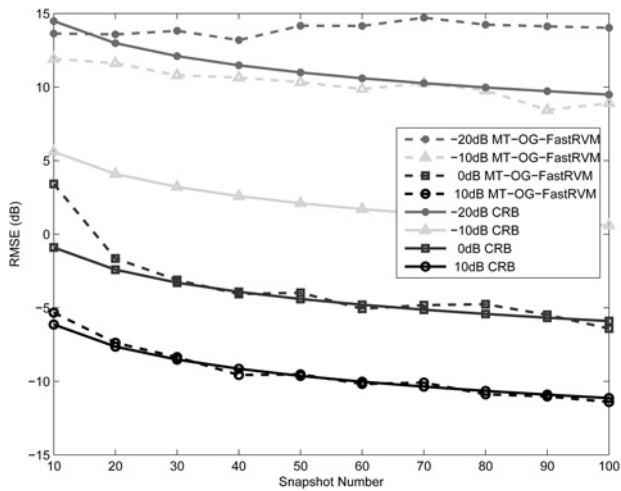


Fig. 6 MT-OG-FastRVM's RMSE and CRB against different numbers of snapshots: $K = 2$, $M = 8$, 200 trials

First, the performances of both FastRVM and OG-FastRVM algorithms are illustrated in Fig. 2. We have done 20 experiments in Fig. 2, and the DOAs of five signals in each experiment are $[-73.7^\circ, -30.32^\circ, -25.68^\circ, -2^\circ, 43.4^\circ]$. In this simulation, the SNR is 20 dB. Fig. 2a shows the overview of 20

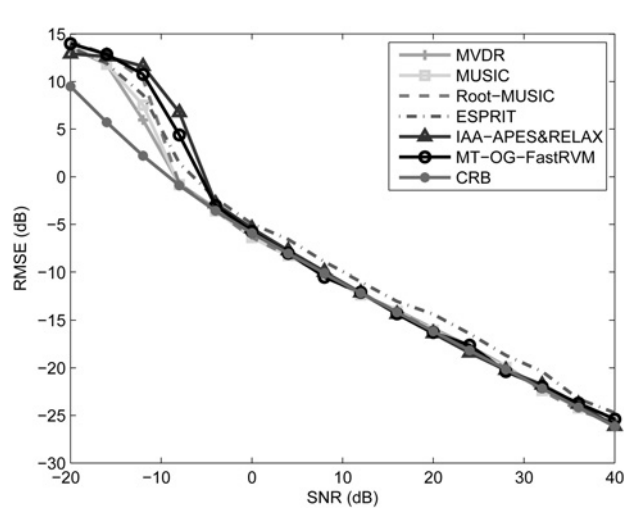


Fig. 9 RMSEs of different algorithms for DOA estimation and CRB against SNR: $T = 100$, $K = 2$, $M = 8$, 200 trials

experiments for the estimation of five DOAs. The amplitudes and directions of the simulation signals denoted by circle are shown in Fig. 2a, and Fig. 2b shows some details of Fig. 2a. It is evident that OG-FastRVM can get a more accurate estimation than FastRVM without the OG technique.

Fig. 3 shows the computational efficiencies of the four single-snapshot algorithms, which are ℓ_1 -penalisation, OG-SBI, IAA-APES&RELAX and OG-FastRVM (All numerical experiments are carried out in MATLAB v.8.3.0 on a personal computer with a 2.67 GHz central processing unit and a Windows

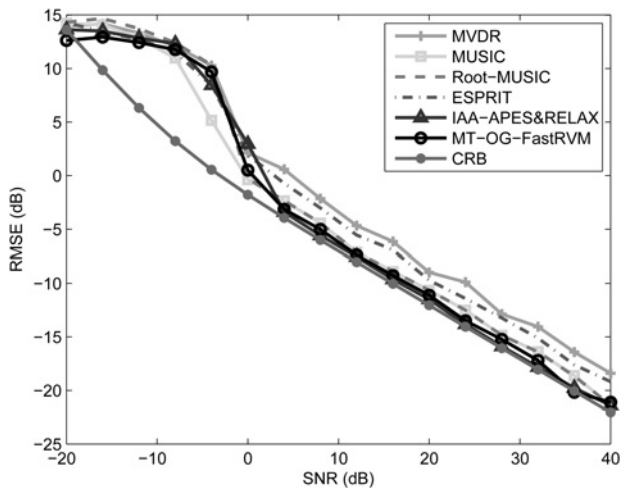


Fig. 7 RMSEs of different algorithms for DOA estimation and CRB against SNR: $T = 10$, $K = 2$, $M = 8$, 200 trials

- 1 $i := \arg \max_i \{ \frac{\|\phi_i^T \mathbf{y}\|^2}{\|\phi_i\|^2} \};$
- 2 select $\phi_i;$
- 3 $\alpha_i := \|\phi_i\|^2 / (\|\phi_i^T \mathbf{t}\|^2 / \|\phi_i\|^2 - \alpha_0^{-1});$
- 4 $\boldsymbol{\mu} := \alpha_0 \boldsymbol{\Sigma} \boldsymbol{\Phi}^T \mathbf{y};$
- 5 $\boldsymbol{\Sigma} := (\boldsymbol{\Lambda} + \alpha_0 \boldsymbol{\Phi}^T \boldsymbol{\Phi})^{-1};$
- 6 $\mathbf{D} \triangleq \alpha_0 \mathbf{I};$
- 7 for $n = 1 : 2N$ do
- 8 $S_n := \phi_n^T \mathbf{D} \phi_n - \phi_n^T \mathbf{D} \boldsymbol{\Phi} \boldsymbol{\Sigma} \boldsymbol{\Phi}^T \mathbf{D} \phi_n;$
- 9 $Q_n := \phi_n^T \mathbf{D} \mathbf{y} - \phi_n^T \mathbf{D} \boldsymbol{\Phi} \boldsymbol{\Sigma} \boldsymbol{\Phi}^T \mathbf{D} \mathbf{y};$
- 10 $s_n := \frac{\alpha_n S_n}{\alpha_n - S_n};$
- 11 $q_n := \frac{\alpha_n Q_n}{\alpha_n - S_n};$

Fig. 10 Initialisation of fast RVM

```

1  $\alpha_0 \triangleq \sigma^{-2}$ ;
2  $\mathbf{e}_i \triangleq \phi_i - \alpha_0 \Phi \Sigma \Phi^T \phi_i$ ;
3  $\alpha_i^{new} := s_i^2 / \theta_i$ ;
4  $\Sigma_{ii} := (\alpha_i^{new} + S_i)^{-1}$ ;
5  $\mu_i^{new} := \Sigma_{ii} Q_i$ ;
6  $\boldsymbol{\mu}^{new} := \begin{bmatrix} \boldsymbol{\mu} - \mu_i \alpha_0 \Sigma \Phi^T \phi_i \\ \mu_i \end{bmatrix}$ ;
7  $\Sigma^{new} := \begin{bmatrix} \Sigma + \alpha_0^2 \Sigma_{ii} \Sigma \Phi^T \phi_i \phi_i^T \Phi \Sigma & -\alpha_0^2 \Sigma_{ii} \Sigma \Phi^T \phi_i \\ -\alpha_0^2 \Sigma_{ii} (\Sigma \Phi^T \phi_i)^T & \Sigma_{ii} \end{bmatrix}$ ;
8  $S_m^{new} := S_m - \Sigma_{ii} (\alpha_0 \phi_i^T \mathbf{e}_i)^2$ ;
9  $Q_m^{new} := Q_m - \mu_i (\alpha_0 \phi_i^T \mathbf{e}_i)$ ;
10  $\Delta \mathcal{L} := \frac{1}{2} \left( \frac{Q_i^2 - S_i}{S_i} + \log \frac{S_i}{Q_i} \right)$ ; /* the increased ML value */

```

Fig. 11 Adding a new basis ϕ_i for fast RVM

```

1  $\alpha_0 \triangleq \sigma^{-2}$ ;
2  $\alpha_i^{new} := s_i^2 / \theta_i$ ;
3  $\kappa_j \triangleq (\Sigma_{jj} + (\alpha_i^{new} - \alpha_i)^{-1})^{-1}$ ;
4  $\Sigma_j :=$  the  $j$ -th column of  $\Sigma$ ; /*  $j$  corresponds to  $i$  */
5  $\boldsymbol{\mu}^{new} := \boldsymbol{\mu} - \kappa_j \mu_j \Sigma_j$ ;
6  $\Sigma^{new} := \Sigma - \kappa_j \Sigma_j \Sigma_j^T$ ;
7  $S_m^{new} := S_m + \kappa_j (\alpha_0 \Sigma_j^T \Phi^T \phi_m)^2$ ;
8  $Q_m^{new} := Q_m + \kappa_j \mu_j (\alpha_0 \Sigma_j^T \Phi^T \phi_m)$ ;
9  $\Delta \mathcal{L} := \frac{1}{2} \left( \frac{Q_i^2}{S_i + [(\alpha_i^{new})^{-1} - \alpha_i^{-1}]^{-1}} - \log \{ 1 + S_i [(\alpha_i^{new})^{-1} - \alpha_i^{-1}] \} \right)$ ; /* the increased ML value */

```

Fig. 12 Re-estimate parameters for fast RVM

7 system.). When the sampling grid is sparse, OG-SBI has a more efficient computation than ℓ_1 -penalisation; when the sampling grid is dense, OG-SBI is slower than ℓ_1 -penalisation. This issue has been pointed out in [31]. Theoretically, in the single-snapshot scenario, the computational complexities of ℓ_1 -penalisation [11], OG-SBI [31] and IAA-APES&RELAX [6] are $\mathcal{O}(N^3)$, $\mathcal{O}(MN^2)$ and $\mathcal{O}(M^2N)$, respectively. The fast RVM algorithm has a computational complexity of $\mathcal{O}(K^2N)$. The complexity of the OG adjustment in OG-FastRVM is negligible compared with that of FastRVM. Actually, our numerical experiments in Fig. 3 have confirmed the theoretical computational complexities of different algorithms. Precisely, the computational efficiencies of IAA-APES&RELAX and OG-FastRVM outperform the ones of ℓ_1 -penalisation and OG-SBI, no matter how dense the sampling grid is. Under the assumption of truly sparse signals ($N \gg M > K$), the complexities mainly depend on N , thus the computational efficiency of the OG-FastRVM algorithm approximates that of IAA-APES&RELAX.

In Fig. 4, we compare OG-FastRVM with ℓ_1 -penalisation, OG-SBI and IAA-APES&RELAX in terms of RMSE. In this experiment, we consider an ULA with $M=8$ array elements and half-wavelength inter-element spacing. The scanning grid interval r is set to 2° , 1° and 0.5° , and the SNR is from 0 to 40 dB. In each trial, the directions of $K=2$ sources are $[-30.3^\circ, -15.3^\circ]$. For each point in Fig. 4, the RMSE is calculated based on $R=200$ trials:

$\text{RMSE} = \sqrt{(1/KR) \sum_{k=1}^K \sum_{i=1}^R (\theta_k - \hat{\theta}_k^i)^2}$, where the superscript $(\cdot)^i$ means the i th trial. It should be noted that the RMSEs of the on-grid methods (ℓ_1 -penalisation and FastRVM) have a lower bound when the SNR is higher. This is because that the best DOA estimates which the on-grid methods can obtain are the grid points closest to the true DOAs. In all conditions of SNR, OG-FastRVM has more accurate DOA estimates than ℓ_1 -penalisation and

```

1  $\bar{\mathbf{B}} = \begin{bmatrix} \mathbf{b}_i^R & -\mathbf{b}_i^I \\ \mathbf{b}_i^I & \mathbf{b}_i^R \end{bmatrix}$ ;
2  $\bar{\boldsymbol{\mu}}_i = [\tilde{\mu}_i, \tilde{\mu}_{i+N}]^T = [\mu_i^R, \mu_i^I]^T$ ;
3  $\Delta \beta_i = \frac{(\tilde{\mathbf{y}} - \bar{\mathbf{A}} \bar{\boldsymbol{\mu}})^T \bar{\mathbf{B}} \bar{\boldsymbol{\mu}}_i}{\bar{\boldsymbol{\mu}}_i^T \bar{\mathbf{B}} \bar{\boldsymbol{\mu}}_i}$ ;
4 if  $\beta_i + \Delta \beta_i \in [-\frac{1}{2}r, \frac{1}{2}r]$  then
5  $\beta_i^{new} = \beta_i + \Delta \beta_i$ ;

```

Fig. 13 Calculate the adjustment factor β_i

```

1  $\alpha_0 \triangleq \sigma^{-2}$ ;
2  $\Sigma_j :=$  the  $j$ -th column of  $\Sigma$ ; /*  $j$  corresponds to  $i$  */
3  $\boldsymbol{\mu}^{new} := \boldsymbol{\mu} - \frac{\mu_j}{\Sigma_{jj}} \Sigma_j$ ;
4  $\Sigma^{new} := \Sigma - \frac{1}{\Sigma_{jj}} \Sigma_j \Sigma_j^T$ ;
5  $S_m^{new} := S_m + \frac{1}{\Sigma_{jj}} (\alpha_0 \Sigma_j^T \Phi^T \phi_m)^2$ ;
6  $Q_m^{new} := Q_m + \frac{\mu_j}{\Sigma_{jj}} (\alpha_0 \Sigma_j^T \Phi^T \phi_m)$ ;
7  $\Delta \mathcal{L} := \frac{1}{2} \left[ \frac{Q_j^2}{S_j - \alpha_j} - \log \left( 1 - \frac{S_j}{\alpha_j} \right) \right]$ ; /* the increased ML value */

```

Fig. 14 Deleting a basis ϕ_i for fast RVM

FastRVM. The OG-FastRVM algorithm has a smaller RMSE than the lower bound caused by the fixed scanning grid. The other two OG methods, OG-SBI and IAA-APES&RELAX can also break through the grid interval limitation and have the similar performances. In most cases of these experiments, OG-FastRVM has a smaller RMSE than the other algorithms. From Figs. 3 and 4, we can observe that IAA-APES&RELAX and our proposed OG-FastRVM method have the advantages of both computational efficiency and RMSE in comparison with the other state-of-the-art single-snapshot methods. From Fig. 4, it can be observed that the RMSEs of OG-FastRVM with different values of r do not keep decreasing with the increasing SNR. It might be that when the SNR is higher, the adjustment factor β_i , $i=1, \dots, N$ cannot have an adequate value, because it depends on $\bar{\boldsymbol{\mu}}$ in (26). Our possible future research tracks might include the further investigation for the influence by the OG technique to DOA estimation. However, in most cases of this numerical simulation, OG-FastRVM has the advantages of both computational complexity and estimation accuracy.

5.2 MT OG-FastRVM

Dealing with multiple snapshots, we have developed the multi-task (MT) implementation of the OG-FastRVM estimator. In this section, the performances of MT-OG-FastRVM are compared with the ones of OG-FastRVM and other classical DOA estimation methods against varying SNRs and snapshot numbers. The theoretical part refers to Section 4.2.

Fig. 5 shows the RMSE values of OG-FastRVM and MT-OG-FastRVM versus varying SNRs, with an array of $M=8$ elements. The performance of MT-OG-FastRVM with variant snapshot numbers has also been illustrated. As it can be observed, the MT-OG-FastRVM method outperforms the single-snapshot FastRVM and OG-FastRVM methods. With the increase of the number of snapshots ($T=10, 20$ and 100), the RMSE performance of MT-OG-FastRVM improves.

These conclusions are further confirmed by the results in Fig. 6 concerned with different SNRs (SNR = -20, -10, 0 and 10 dB). Moreover, the Cramer-Rao bounds (CRBs) [2, 4] of different SNRs versus variant snapshot numbers are added in Fig. 6 for comparison. It can be observed that the larger the number of snapshots, the lower is the RMSE of the MT-OG-FastRVM method. When the SNR is larger than a specific value, the RMSE approaches the CRB, and the following simulations show the details.

The final numerical simulation is concerned with a comparative assessment of MT-OG-FastRVM and the state-of-the-art algorithms such as MUSIC, Root-MUSIC, ESPRIT and IAA-APES&RELAX. The CRB is also added in these

comparisons. Figs. 7–9 plot the RMSEs averaged over 200 trials for each point and show the different performances with variant snapshot numbers ($T=10, 20$ and 100). It can be seen that the RMSE of MT-OG-FastRVM follows closely those of other estimators and, except for very low SNR, meets the CRB. Especially for a small number of snapshots ($T=10$), IAA-APES&RELAX and our proposed MT-OG-FastRVM algorithm outperform the others.

6 Conclusion

In this paper, an algorithm termed OG-FastRVM has been proposed to estimating the DOAs from a single snapshot. The proposed algorithm is based on the fast RVM, and utilises the OG technique in each iteration. The results of numerical simulations have illustrated that the proposed approach outperforms the standard ℓ_1 -penalisation algorithm. Furthermore, OG-FastRVM computes more efficiently than both ℓ_1 -penalisation and OG-SBI.

Furthermore, the MT-OG-FastRVM has also been realised by the combination of OG-FastRVM and MT-BCS. The DOA estimation performance of this MT method exceeds the single-snapshot (OG-) FastRVM. Precisely, with the increase of the number of snapshots, the RMSE decreases. Finally, we compare the MT-OG-FastRVM method with the state-of-the-art ones. Especially with a small number of snapshots, the proposed method provides more accurate estimates than the other compared methods.

7 Acknowledgments

We thank the Editor Hugh Griffiths and two anonymous referees who provided insightful comments and constructive criticisms that greatly improved the manuscript. This work was supported by the National Natural Science Foundation of China under grant numbers 61222107 and 61471352.

8 References

- 1 Krim, H., Viberg, M.: 'Two decades of array signal processing research: the parametric approach', *IEEE Signal Process. Mag.*, 1996, **13**, (4), pp. 67–94
- 2 Van Trees, H.L.: 'Optimum array processing (detection, estimation, and modulation theory, Part IV)' (Wiley, New York, 2002)
- 3 Schmidt, R.O.: 'Multiple emitter location and signal parameter estimation', *IEEE Trans. Antennas Propag.*, 1986, **34**, (3), pp. 276–280
- 4 Stoica, P., Arye, N.: 'Music, maximum likelihood, and Cramer–Rao bound', *IEEE Trans. Acoust. Speech Signal Process.*, 1989, **37**, (5), pp. 720–741
- 5 Roy, R., Kailath, T.: 'ESPRIT-estimation of signal parameters via rotational invariance techniques', *IEEE Trans. Acoust. Speech Signal Process.*, 1989, **37**, (7), pp. 984–995
- 6 Yardibi, T., Li, J., Stoica, P., et al.: 'Source localization and sensing: a nonparametric iterative adaptive approach based on weighted least squares', *IEEE Trans. Aerosp. Electron. Syst.*, 2010, **46**, (1), pp. 425–443
- 7 Roberts, W., Stoica, P., Li, J., et al.: 'Iterative adaptive approaches to MIMO radar imaging', *IEEE J. Sel. Top. Signal Process.*, 2010, **4**, (1), pp. 5–20
- 8 Yardibi, T., Li, J., Stoica, P.: 'Nonparametric and sparse signal representations in array processing via iterative adaptive approaches'. 2008 42nd Asilomar Conf. on Signals, Systems and Computers, 2008, pp. 278–282
- 9 Stoica, P., Li, J., Ling, J.: 'Missing data recovery via a nonparametric iterative adaptive approach', *IEEE Signal Process. Lett.*, 2009, **16**, (4), pp. 241–244
- 10 Fortunati, S., Grasso, R., Gini, F., et al.: 'Single-snapshot DOA estimation by using compressed sensing', *EURASIP J. Adv. Signal Process.*, 2014, (1), pp. 1–17
- 11 Malioutov, D., Çetin, M., Willsky, A.: 'A sparse signal reconstruction perspective for source localization with sensor arrays', *IEEE Trans. Signal Process.*, 2005, **53**, (8), pp. 3010–3022
- 12 Li, X., Ma, X., Yan, S., et al.: 'Single snapshot DOA estimation by compressive sampling', *Appl. Acoust.*, 2013, **74**, (7), pp. 926–930
- 13 Stoica, P., Selen, Y.: 'Model-order selection: a review of information criterion rules', *IEEE Signal Process. Mag.*, 2004, **21**, (4), pp. 36–47
- 14 Li, J., Stoica, P.: 'Efficient mixed-spectrum estimation with applications to target feature extraction', *IEEE Trans. Signal Process.*, 1996, **44**, (2), pp. 281–295

- 15 Li, J., Stoica, P.: 'Angle and waveform estimation via RELAX', *IEEE Trans. Aerosp. Electron. Syst.*, 1997, **33**, (3), pp. 1077–1087
- 16 Baraniuk, R.: 'Compressive sensing', *IEEE Signal Process. Mag.*, 2007, **24**, (4), pp. 118–121
- 17 Tipping, M.E.: 'Sparse Bayesian learning and the relevance vector machine', *J. Mach. Learn. Res.*, 2001, **1**, (3), pp. 211–244
- 18 Ji, S., Xue, Y., Carin, L.: 'Bayesian compressive sensing', *IEEE Trans. Signal Process.*, 2008, **56**, (6), pp. 2346–2356
- 19 Wipf, D.P., Rao, B.D.: ' ℓ_q -norm minimization for basis selection'. Advances in Neural Information Processing Systems, 2004, pp. 1513–1520
- 20 Rao, B.D., Wipf, D.P.: 'Comparing the effects of different weight distributions on finding sparse representations'. Advances in Neural Information Processing Systems, 2005, pp. 1521–1528
- 21 Carlin, M., Rocca, P., Oliveri, G.: 'Directions-of-arrival estimation through Bayesian compressive sensing strategies', *IEEE Trans. Antennas Propag.*, 2013, **61**, (7), pp. 3828–3838
- 22 Ji, S., Dunson, D., Carin, L.: 'Multi-task compressive sensing', *IEEE Trans. Signal Process.*, 2008, **57**, (1), pp. 92–106
- 23 Liu, Z.M., Huang, Z.T., Zhou, Y.Y.: 'Sparsity-inducing direction finding for narrowband and wideband signals based on array covariance vectors', *IEEE Trans. Wirel. Commun.*, 2013, **12**, (8), pp. 1–12
- 24 Liu, Z.M., Huang, Z.T., Zhou, Y.Y.: 'An efficient maximum likelihood method for direction-of-arrival estimation via sparse Bayesian learning', *IEEE Trans. Wirel. Commun.*, 2012, **11**, (10), pp. 1–11
- 25 Liu, Z.M., Liu, Z., Feng, D.W., et al.: 'Direction-of-arrival estimation for coherent sources via sparse Bayesian learning', *Int. J. Antennas Propag.*, 2014, (2014), pp. 1–8
- 26 Carlin, M., Rocca, P.: 'A Bayesian compressive sensing strategy for direction-of-arrival estimation'. 2012 Sixth European Conf. on Antennas and Propagation (EUCAP), 2012, pp. 1508–1509
- 27 Zhao, G., Wang, Q., Shen, F., et al.: 'Robust compressive multi-input–multi-output imaging', *IET Radar Sonar Navig.*, 2013, **7**, (3), pp. 233–245
- 28 Carlin, M., Rocca, P., Oliveri, G., et al.: 'Probabilistic direction of arrival estimation through Bayesian compressive sensing'. 2014 Eighth European Conf. on IEEE Antennas and Propagation (EuCAP), 2014, pp. 359–362
- 29 Carlin, M., Rocca, P., Oliveri, G., et al.: 'Bayesian compressive sensing as applied to directions-of-arrival estimation in planar arrays', *J. Electr. Comput. Eng.*, 2013, (2013), pp. 1–12
- 30 Wang, Y., Duan, X., Tian, D., et al.: 'A Bayesian compressive sensing vehicular location method based on three-dimensional radio frequency', *Int. J. Distrib. Sens. Netw.*, 2014, (2014), pp. 1–13
- 31 Yang, Z., Xie, L., Zhang, C.: 'Off-grid direction of arrival estimation using sparse Bayesian inference', *IEEE Trans. Signal Process.*, 2013, **61**, (1), pp. 38–43
- 32 Faul, A.C., Tipping, M.E.: 'Analysis of sparse Bayesian learning', *Adv. Neural Inf. Process. Syst.*, 2001, **20**, (3), pp. 383–389
- 33 Tipping, M.E., Faul, A.: 'Fast marginal likelihood maximisation for sparse Bayesian models'. Proc. of the Ninth Int. Workshop on Artificial Intelligence and Statistics, 2003, pp. 3–6
- 34 Zhu, H., Leus, G., Giannakis, G.: 'Sparsity-cognizant total least-squares for perturbed compressive sampling', *IEEE Trans. Signal Process.*, 2011, **59**, (5), pp. 2002–2016
- 35 Tang, G., Bhaskar, B.N., Shah, P., et al.: 'Compressed sensing off the grid', *IEEE Trans. Inf. Theory*, 2013, **59**, (11), pp. 7465–7490
- 36 Tan, Z., Nehorai, A.: 'Sparse direction of arrival estimation using co-prime arrays with off-grid targets', *IEEE Signal Process. Lett.*, 2014, **21**, (1), pp. 26–29
- 37 Ozay, M., Esnaola, I., Vural, F.T., et al.: 'Sparse attack construction and state estimation in the smart grid: centralized and distributed models', *IEEE J. Sel. Areas Commun.*, 2013, **31**, (7), pp. 1306–1318
- 38 Yang, Z., Xie, L.: 'On gridless sparse methods for line spectral estimation from complete and incomplete data', arXiv preprint arXiv:1407.2490, 2014
- 39 Donoho, D.L.: 'Compressed sensing', *IEEE Trans. Inf. Theory*, 2006, **52**, pp. 1289–1306
- 40 Goodman, N.: 'Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction)', *Ann. Math. Stat.*, 1963, **34**, (1), pp. 152–177
- 41 Oliveri, G., Massa, A.: 'Bayesian compressive sampling for pattern synthesis with maximally sparse non-uniform linear arrays', *IEEE Trans. Antennas Propag.*, 2011, **59**, (2), pp. 467–481

9 Appendix

The fast RVM algorithm proposed in [33] is customised to deal with the OG model for DOA estimation. The details are presented in Algorithms 2–6 (Figs. 10–14) (The number of the basis functions at step t of the OG-FastRVM is denoted as N_t , $i \in \{1, \dots, N\}$ indexes the new single basis function not existing in the current basis, $j \in \{1, \dots, N_t\}$ indexes the single basis function and associated parameters that already exist in the current basis.). Here, for simplification, $\tilde{\Phi}$, $\tilde{\mathbf{y}}$, $\tilde{\boldsymbol{\mu}}$, $\tilde{\boldsymbol{\alpha}}$ and σ^{-2} are represented as Φ , \mathbf{y} , $\boldsymbol{\mu}$, $\boldsymbol{\alpha}$ and α_0 , respectively.