

Communication

Wideband DOA Estimation Utilizing Hierarchical Prior Based on Variational Bayesian Inference

Ninghui Li ^{1,*}, Xiaokuan Zhang ², Binfeng Zong ², Fan Lv ¹, Jiahua Xu ¹ and Zhaolong Wang ¹

¹ Graduate school, Air Force Engineering University, Xi'an, 710051, China

² Air Defence and Anti-missile Institute and school, Air Force Engineering University, Xi'an, 710051, China

* Correspondence: li6667778882022@163.com

Abstract: Sparse direction-of-arrival (DOA) estimation of wideband signal has attracted widespread researches for its unique high-resolution performance. Numerous existing methods based on sparse Bayesian learning (SBL) don't possess the ability to enhance sparsity even if they have already enjoyed high favor among sparse recovery approaches. In view of this, we propose a novel hierarchical Bayesian prior framework that enhance sparsity evidently and derive its corresponding iterative algorithm. On analysis, the computational complexity of the iterative are less than most existing other state-of-the-art algorithms. Not only that, proposed method possesses high angular estimation precision and sparsity performance by utilizing joint sparsity of multiple measure vector (MMV) models. Last but not least, the method obtains the ability to stabilize the estimated values between different frequencies or snapshots such that flat spatial spectrum. Extensive simulation results are presented to prove the superior performance of our methods

Keywords: direction-of-arrival; sparse Bayesian learning; hierarchical Bayesian prior; sparse recovery;

1. Introduction

DOA estimation is an important research focus that has developed rapidly during the past three decades in array signal processing, and its applications are spread across military and civilian fields such as radar, sonar, wireless detection, mobile communication, and biomedicine [1]. It is a remarkable fact that most current researches are based on narrowband signal, while the study for wideband-signal DOA estimation is of equal importance since the common assumption that signals incident from different angles occupy the same frequency band does not always hold in practice [2]-[3]. In addition, compared to narrowband signals, band signals are characterized by high resolution, which means higher estimation accuracy and better robustness to correlated sources.

Traditional DOA estimation methods for wideband signals can be roughly divided into two categories, one is incoherent signal subspace methods (ISSM) [4]-[5] and the other is coherent signal subspace methods (CSSM) [6]-[8]. The common point between them is to convert no-spectrum-aliasing wideband signals into narrowband signals for processing, while the most obvious difference is that CSSM can handle coherent sources with lower operational complexity. It is worth noting that there are two methods based on orthogonal testing between ISSM and CSSM, though some literature classifies them as special ISSM category. One is the test of orthogonality of projected subspaces (TOPS), the other is the test of orthogonality of frequency subspaces (TOFS). TOPS and TOFS are with different principles such that they place different emphasis on DOA estimation problems. Their common features are that they can solve the problem of seeking the optimal reference frequency point, and find a balance between the signal to noise ratio (SNR) and the computational complexity. However, all of them need to face the following basic limitations: (i) Most of them need pre-estimation procedure. (ii) The relatively higher snapshot accumulation is necessary for accurate angular estimation. (iii) A priori of the number of sources is required. (iv) Most of their performance is constrained by the SNR.

For aforementioned limitations, many algorithms based on compressed sensing and sparse representation have been proposed and gradually developed [9]-[10]. These methods possess better performance than traditional ones such as locating coherent sources well, depending less on high SNR and snapshots and operating with higher efficiency. According to the current research, the methods based on sparse representation for wideband-signal DOA estimation can be broadly classified into three categories: the first one is based on basis pursuit, such as matching pursuit (MP) [11] and orthogonal matching pursuit (OMP) [12]; the second one is based on convex optimization, such as JLZA [13], l_1 -SVD [14] and W-SpSF [15]; the final one is based on Bayesian compressed sensing [16]-[17]. The methods' comprehensive performance of sparse Bayesian learning (SBL) is reported to be better than that based on basis pursuit or convex optimization, though this conclusion was drawn by narrowband signals [18]. There exist several methods based on sparse Bayesian learning, but they all adopted single measure vector model (SMV) [17] or generalized SMV model [16]. In other words, these methods aim at single-snapshot models or vectorized multiple-snapshot models, which would result in a simple decoupling between snapshots. In fact, a reasonable use of temporal correlation not only represents the time-domain coherence of sources, but also improves the algorithm performance. In conclusion, it has been shown that the MMV model has better properties than the SMV model (e.g., stronger joint sparsity and better sparse recovery) [19]-[20].

In view of this, a novel method, based on the block sparse Bayesian model for wideband-signal DOA estimation, is developed in this paper. The method can be considered as one extension of the SBL based on the SMV model to the MMV model, or as the other extension of the SBL based on narrowband DOA estimation to wideband-signal one. On the basis of SBL, the method takes full advantage of the MMV model by transforming the MMV model into block-sparse model. Not only that, the tactfully design for proposed algorithm allows the joint sparsity between snapshot and spectrum to be exploited simultaneously. In addition, compared to the traditional Gaussian prior, the employed hierarchical structure of the prior can greatly improve the induced sparsity of sparse Bayesian learning.

The rest of this work is organized as follows. In Section 2, the data model is introduced and transformed. In Section 3, the likelihood and prior are presented for variational Bayesian inference and computational complexity of various approaches are compared. The performance of our algorithms is evaluated in Section 4 and the conclusions are drawn in Section 5.

2. Data model

Assume K uncorrelated far-field wideband sources from different directions $\boldsymbol{\theta} = \{\theta_1, \dots, \theta_K\}$ impinge on a linear array with N sensors. Without loss of generality, corresponding received data is modeled as

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

where $\mathbf{A} = [\mathbf{a}(\theta_1), \dots, \mathbf{a}(\theta_K)] \in \mathbb{C}^{N \times K}$ is array manifold matrix with its k -th entry is $\mathbf{a}(\theta_k) = [\exp(-jv_1), \dots, \exp(-jv_N)]^T$, $v_n = 2\pi f d_n \sin(\theta_k) / c$, $n = 1, \dots, N$. d_n denotes the location of the n -th sensor relative to reference one, c is propagation velocity. $\mathbf{s}(t) = [s_1(t), \dots, s_K(t)]$ and $\mathbf{n}(t) = [n_1(t), \dots, n_N(t)]$ are signal waveform vector and additive white Gaussian noise (AWGN) vector, respectively. Then the time-domain data could be converted into frequency-domain data by filter bank (FB) or discrete Fourier transform (DFT).

$$\mathbf{x}(f_j) = \mathbf{A}(f_j)\mathbf{s}(f_j) + \mathbf{n}(f_j) \quad (2)$$

where f_j emphasize that wideband signal is divided into J sub-bands (frequency bins) and their individual center frequency is f_j , $j = 1, \dots, J$. In order to ensure the data's independence in the same sub-band between different snapshots, the total observed time satisfies $T \gg 1/(BL)$, where B is the bandwidth and L is the number of snapshots. On the basis of this, the frequency-domain model of wideband array signal is entirely the same to the time-domain model of narrowband array signal for

model structure. In other words, frequency-domain data can be directly treated as time-domain data. After matched filter and snapshots accumulation, (2) could be casted to

$$\mathbf{X}_j = \mathbf{A}_j \mathbf{S}_j + \mathbf{N}_j \quad (3)$$

where $\mathbf{A}_j = \mathbf{A}(f_j)$, $\mathbf{S}_j \in \mathbb{C}^{K \times L}$ denotes the complex amplitude matrix for L snapshots. $\mathbf{N}_j \in \mathbb{C}^{N \times L}$ is the noise matrix. In order to let the model be available for sparse-recovery methods, (3) needs to be converted to following sparse form by sparse representation.

$$\mathbf{X}_j = \bar{\mathbf{A}}_j \mathbf{P}_j + \mathbf{N}_j \quad (4)$$

where $\bar{\mathbf{A}}_j = [\mathbf{a}_j(\bar{\theta}_1), \dots, \mathbf{a}_j(\bar{\theta}_M)] \in \mathbb{C}^{N \times M}$ is the extended manifold matrix, $\bar{\theta}_m \in \{\bar{\theta}_{m=1}^M\}$, which is produced by discrete sampling. $\mathbf{P} \in \mathbb{C}^{M \times L}$ is solution matrix and its every row represents a potential source. Therefore, the problem of DOA estimation is casted sparse recovery problem of (4). In terms of [21], (4) can be rewritten as follows.

$$\mathbf{y}_j = \Phi_j \mathbf{p}_j + \mathbf{n}_j \quad (5)$$

where $\mathbf{y}_j = \text{vec}(\mathbf{X}_j^T) \in \mathbb{C}^{NL \times 1}$, $\Phi_j = \bar{\mathbf{A}}_j \otimes \mathbf{I}_L \in \mathbb{C}^{NL \times ML}$, $\mathbf{p}_j = \text{vec}(\mathbf{P}_j^T) \in \mathbb{C}^{ML \times 1}$, $\mathbf{n}_j = \text{vec}(\mathbf{N}_j^T) \in \mathbb{C}^{NL \times 1}$. (5) is block-sparse model since \mathbf{p}_j is segmentally sparse (i.e., many segments only contain zeros). As for \mathbf{p}_j , we just care about the locations of nonzero elements rather than the concrete values, and different \mathbf{p}_j indicate the same locations of sources. Therefore, different \mathbf{p}_j can be unified as \mathbf{p} . Consider all the frequency bins, (5) could be extended as

$$\mathbf{y} = \Psi \mathbf{p} + \bar{\mathbf{n}} \quad (6)$$

$$\text{where } \bar{\mathbf{y}} = [\mathbf{y}_1^T, \dots, \mathbf{y}_J^T]^T \in \mathbb{C}^{NLJ \times 1}, \quad \Psi = [\Phi_1^T, \dots, \Phi_J^T]^T \in \mathbb{C}^{NLJ \times ML}, \\ \bar{\mathbf{n}} = [\mathbf{n}_1^T, \dots, \mathbf{n}_J^T]^T \in \mathbb{C}^{NLJ \times 1}.$$

3. Proposed approach

3.1. Bayesian Model

It is necessary and vital to construct appropriate Bayesian model for sparse Bayesian learning. Priors with hierarchy structure are imposed on hidden variables so as to favouring sparsity more, and the details are as below. As for observed variable \mathbf{y} , the likelihood is

$$p(\mathbf{y} | \mathbf{p}; \delta) \sim \mathcal{CN}(\Psi \mathbf{p}, \delta^{-1} \mathbf{I}_{NLJ}) \quad (7)$$

Impose Gaussian distribution prior on hidden variable \mathbf{p} such that

$$p(\mathbf{p}; \boldsymbol{\gamma}) \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Sigma}) \quad (8)$$

where $\boldsymbol{\gamma} = [\gamma_1^{-1}, \dots, \gamma_M^{-1}]^T$, $\boldsymbol{\Sigma} = \text{diag}(\boldsymbol{\gamma}) \otimes \mathbf{I}_L \in \mathbb{R}^{ML \times 1}$. Since inverse Gamma distribution is conjugate to Gaussian distribution, the Gamma distribution prior is adopted for each element of $\boldsymbol{\gamma}$.

$$p(\boldsymbol{\gamma}; a, b) = \prod_{m=1}^M \frac{b^a}{\Gamma(a)} \gamma_m^{a-1} e^{-b\gamma_m} \quad (9)$$

where $\Gamma(a) = \int_0^\infty x^{a-1} \exp(-x) dx$. a is the shape parameter and b is the scale parameter. Similarly, assume δ obeys Gamma distribution, which yields

$$p(\delta; c, d) = \frac{c^d}{\Gamma(c)} \delta^{c-1} e^{-d\delta} \quad (10)$$

where c and d are corresponding shape and scale parameters, respectively. The directed acyclic graph for representing Bayesian model is as shown in Figure 1.

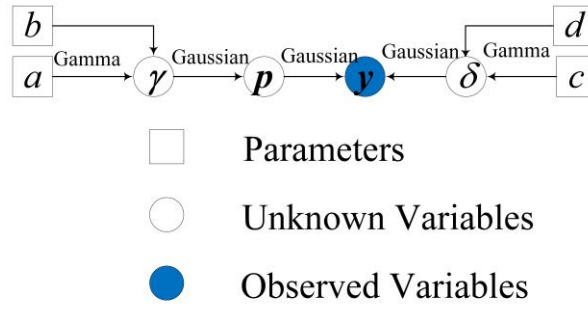


Figure 1. Directed acyclic graph of proposed Bayesian model.

3.2. Variational Bayesian inference

In the purpose of deriving the iterative algorithm based on SBL, Bayesian inference is adopted. Unluckily, the closed-form solution for posterior cannot be directly obtained due to limitations of mathematics. But the approximate solution would be gained by variational Bayesian inference [21]. The posterior is factorized as below.

$$p(\mathbf{p}, \boldsymbol{\gamma}, \boldsymbol{\delta} | \mathbf{y}; a, b, c, d) \approx q(\mathbf{p}, \boldsymbol{\gamma}, \boldsymbol{\delta}) = q(\mathbf{p})q(\boldsymbol{\gamma})q(\boldsymbol{\delta}) \quad (11)$$

where $q(\mathbf{p})$, $q(\boldsymbol{\gamma})$ and $q(\boldsymbol{\delta})$ are separable marginal distributions of \mathbf{p} , $\boldsymbol{\gamma}$ and $\boldsymbol{\delta}$. Each of their logarithmic forms can be solved by the others. As for $\ln q(\mathbf{p})$, it satisfies

$$\ln q(\mathbf{p}) = \langle \ln p(\mathbf{y} | \mathbf{p}, \boldsymbol{\delta}) p(\mathbf{p}; \boldsymbol{\gamma}) \rangle_{q(\boldsymbol{\gamma})q(\boldsymbol{\delta})} + \text{const} \quad (12)$$

In terms of (7), (8) and (12), $q(\mathbf{p})$ can be solved to obey the Gaussian distribution, whose mean and variance are as below.

$$\boldsymbol{\mu}_p = \langle \boldsymbol{\delta} \rangle \boldsymbol{\Sigma}_p \boldsymbol{\Psi}^H \mathbf{y} \quad (13)$$

$$\boldsymbol{\Sigma}_p = (\langle \boldsymbol{\delta} \rangle \boldsymbol{\Psi}^H \boldsymbol{\Psi} + \langle \boldsymbol{\Sigma} \rangle^{-1})^{-1} \quad (14)$$

As usual, $NJ < M$ holds when dense sampling is adopted for high precision estimation, so (13) and (14) could be equivalently transformed into following formulas for less computational complexity.

$$\boldsymbol{\mu}_p = \langle \boldsymbol{\Sigma} \rangle \boldsymbol{\Psi}^H (\langle \boldsymbol{\delta} \rangle^{-1} \mathbf{I}_{NLJ} + \boldsymbol{\Psi} \langle \boldsymbol{\Sigma} \rangle \boldsymbol{\Psi}^H)^{-1} \mathbf{y} \quad (15)$$

$$\boldsymbol{\Sigma}_p = \langle \boldsymbol{\Sigma} \rangle - \langle \boldsymbol{\Sigma} \rangle \boldsymbol{\Psi}^H (\langle \boldsymbol{\delta} \rangle^{-1} \mathbf{I}_{NLJ} + \boldsymbol{\Psi} \langle \boldsymbol{\Sigma} \rangle \boldsymbol{\Psi}^H)^{-1} \boldsymbol{\Psi} \langle \boldsymbol{\Sigma} \rangle \quad (16)$$

Likewise, $q(\boldsymbol{\gamma})$ satisfies

$$\ln q(\boldsymbol{\gamma}) = \langle \ln p(\mathbf{p}; \boldsymbol{\gamma}) p(\boldsymbol{\gamma}; a, b) \rangle_{q(\mathbf{p})q(\boldsymbol{\delta})} + \text{const} \quad (17)$$

Utilizing (8), (9) and (17), $q(\boldsymbol{\gamma})$ is identified as a Gamma distribution, whose shape parameter, m -th scale parameter and mean are as below.

$$\bar{a} = a + \frac{1}{2} \quad (18)$$

$$\bar{b}_m = b + \frac{1}{2} \langle \mathbf{p}_m^H \mathbf{p}_m \rangle = b + \frac{1}{2} \boldsymbol{\mu}_{\mathbf{p}_m}^H (\mathbf{I}_L + \text{diag}(\text{diag}(\boldsymbol{\Sigma}_{\mathbf{p}_m}))) \boldsymbol{\mu}_{\mathbf{p}_m} \quad (19)$$

$$\langle \gamma_m \rangle = \frac{\bar{a}}{\bar{b}_m} \quad (20)$$

where $\mathbf{p}_m, \boldsymbol{\mu}_{\mathbf{p}_m} \in \mathbb{C}^{L \times 1}$ is respectively the m -th entry of $\mathbf{p} = [\mathbf{p}_1^T, \dots, \mathbf{p}_M^T]^T$, $\boldsymbol{\mu}_{\mathbf{p}} = [\boldsymbol{\mu}_{\mathbf{p}_1}^T, \dots, \boldsymbol{\mu}_{\mathbf{p}_M}^T]^T$, $m = 1, 2, \dots, M$. $\boldsymbol{\Sigma}_{\mathbf{p}_m} = \boldsymbol{\Sigma}_{\mathbf{p}}([((m-1)L+1:mL), ((m-1)L+1:mL)]) \in \mathbb{C}^{L \times L}$ is the sub-matrix of $\boldsymbol{\Sigma}_{\mathbf{p}}$ according to MATLAB notations.

Similarly, $q(\delta)$ satisfies

$$\ln q(\delta) = \langle \ln p(\mathbf{y} | \mathbf{p}, \delta) p(\delta; c, d) \rangle_{q(\mathbf{p})q(\gamma)} + \text{const} \quad (21)$$

With (7), (10) and (21), $q(\delta)$ can be also solved as a Gamma distribution and its shape parameter \bar{c} , scale parameter \bar{d} and mean are as below.

$$\bar{c} = c + \frac{NLJ}{2} \quad (22)$$

$$\begin{aligned} \bar{d} &= d + \frac{1}{2} \langle (\mathbf{y} - \boldsymbol{\Psi}\mathbf{p})^H (\mathbf{y} - \boldsymbol{\Psi}\mathbf{p}) \rangle = d + \frac{1}{2} \mathbf{y}^H \mathbf{y} - \text{real}(\boldsymbol{\mu}_{\mathbf{p}}^H \boldsymbol{\Psi}^H \mathbf{y}) + \frac{1}{2} \langle \boldsymbol{\Psi}^H \boldsymbol{\Psi} \mathbf{p} \mathbf{p}^H \rangle \\ \langle \delta \rangle &= \frac{\bar{c}}{\bar{d}} \end{aligned} \quad (24)$$

For solving the final term of (23), $\boldsymbol{\Psi}^H \boldsymbol{\Psi}$ needs to be divided into two parts, i.e., $(\boldsymbol{\Psi}^H \boldsymbol{\Psi})_{\bar{\lambda}} = \text{diag}(\text{diag}(\boldsymbol{\Psi}^H \boldsymbol{\Psi}))$, $(\boldsymbol{\Psi}^H \boldsymbol{\Psi})_{\lambda} = \boldsymbol{\Psi}^H \boldsymbol{\Psi} - \text{diag}(\text{diag}(\boldsymbol{\Psi}^H \boldsymbol{\Psi}))$. Therefore, (23) can be rewritten as

$$\bar{d} = d + \frac{1}{2} \mathbf{y}^H \mathbf{y} - \text{real}(\boldsymbol{\mu}_{\mathbf{p}}^H \boldsymbol{\Psi}^H \mathbf{y}) + \frac{1}{2} \|(\boldsymbol{\Psi}^H \boldsymbol{\Psi})_{\bar{\lambda}} \odot (\boldsymbol{\mu}_{\mathbf{p}}^H \boldsymbol{\mu}_{\mathbf{p}})\|_{1,1} + \frac{1}{2} \|(\boldsymbol{\Psi}^H \boldsymbol{\Psi})_{\lambda} \odot (\boldsymbol{\mu}_{\mathbf{p}}^H \boldsymbol{\mu}_{\mathbf{p}} + \boldsymbol{\Sigma}_{\mathbf{p}})\|_{1,1} \quad (25)$$

So far, the derivation of our proposed iterative algorithm is completed. Utilizing (15), (16), (20) and (24), our iterative algorithm is completed. The specific steps are as follows.

(1) **Initialization.** First iterative number $k = 0$, $a = b = c = d = 10^{-6}$ (ensure uninformative distribution), $\mathbf{p}^{(0)} = (\boldsymbol{\Psi}^H \boldsymbol{\Psi})^{-1} \boldsymbol{\Psi}^H \mathbf{y}$. Preset error tolerance ε and maximal iterative number k_{\max} .

(2) **Repetition.** Compute hyperparameters with (12), (13), (17) and (22). Iterate alternately until convergence. Regard $\boldsymbol{\mu}_{\mathbf{p}}$ as $\mathbf{p}^{(k)}$ and record $k = k + 1$.

(3) **Check.** If $\|(\mathbf{p}^{(k+1)} - \mathbf{p}^{(k)}) / \mathbf{p}^{(k)}\|_2 \leq \varepsilon$ or $k = k_{\max}$ is satisfied, terminate the algorithm. Otherwise, return to (ii).

(4) **Output.** Obtain final $\mathbf{p}^{(k)}$ and calculate corresponding DOA.

3.3. Computational Complexity

The computational complexity of our algorithm is $O(MN^2L^3J^2)$. When generally dense sampling is adopted, the computational complexity is less than $O(N^2M^2)$. The computational complexity of W-SpSF [15], W-SBL [16], l_1 -SVD [14], JLZA [13] is respectively $O(J^3M^3)$, $O(JM^3)$, $O(K^3M^3)$, $O(M^3 + NM^2 + LNM)$. Generally, our proposed algorithm possesses minimum computational complexity because $M \gg N, L, J, K$ holds in many cases. Of course, the complexity of proposed algorithm is little larger than ISSM or CSSM with many limitations.

4. Numerical Simulation

In this section, the superior performance of our proposed method would be verified by following simulation results. Before that, it is necessary to introduce the Root Mean Square Error (RMSE). RMSE is defined as

$$\text{RMSE} = \sqrt{\frac{1}{M_c K} \sum_{m_c=1}^{M_c} \sum_{k=1}^K (\hat{\theta}_{m_c, k} - \theta_k)^2} \quad (26)$$

where M_c is the Monte Carlo number and $\hat{\theta}_{m_c, k}$ is the angle estimation value for the k -th source in the m_c -th trial, $m_c = 1, \dots, M_c$, $k = 1, \dots, K$.

In order to analyze the performance of our proposed method intuitively, ISSM [4], W-SpSF [15], W-SBL [16], l_1 -SVD [14] and JLZA [13] are introduced as comparison. The basic experimental conditions are as follows. $K=3$ uncorrelated sources, number of sensors $N=8$, central frequency $f_0=400\text{Hz}$, bandwidth $BW=200\text{Hz}$, $J=5$, $M_c=300$, grid interval 1° , number of grids $M=180$. Other conditions are indicated in specific experiments.

Experiment 1 tests angular resolution of various algorithms. Other conditions: DOA set $\{-5^\circ, 35^\circ, 60^\circ\}$ (ensure on grid), $\text{SNR}=20\text{dB}$, number of snapshots $T=20$. As shown in Figure. 2, the estimated values of proposed method and ISSM are most precise. Note that proposed method possesses smoother peak values, which expresses small fluctuation with respect to frequency and proves better joint sparsity performance of proposed

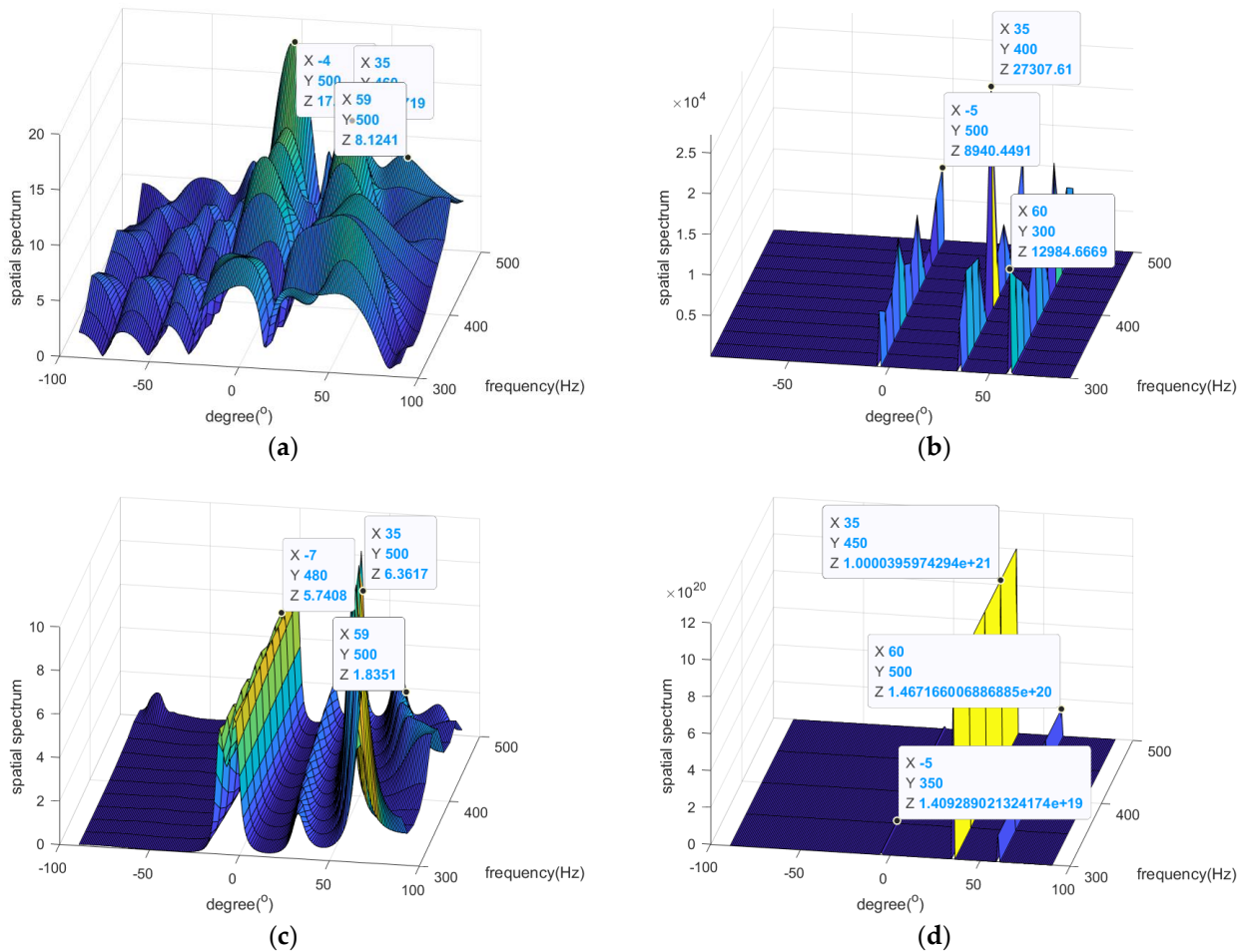


Figure 2. Spatial spectrum versus degree and frequency: (a) JLZA, (b) ISSM, (c) l_1 -SVD, (d) proposed method.

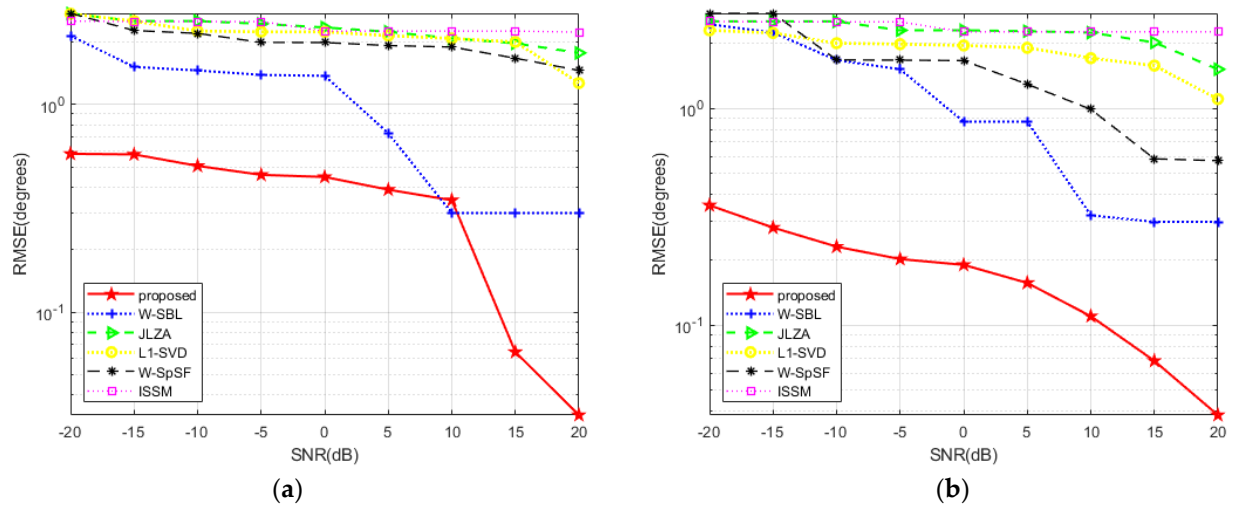


Figure 3. RMSE versus SNR, (a) $T = 5$, (b) $T = 20$.

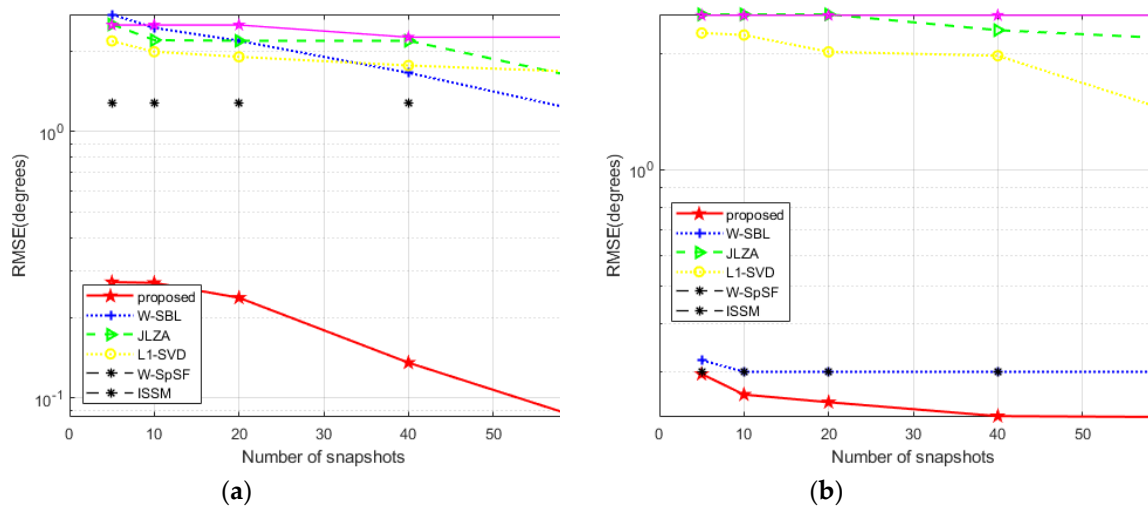


Figure 4. RMSE versus number of snapshots, (a) $SNR = -10dB$, (b) $SNR = 20dB$.

Method than ISSM. The results can be explained by the fact that SBL is able to convergence at sparse solutions and possesses robustness to ordinary condition change itself, and proposed method has the ability to enhance sparsity more and stabilize the estimated values of DOA between different frequencies. In a word, proposed method remains the advantages of sparse Bayesian learning and shows better estimation performance than ordinary methods based on SBL.

Experiment 2 examines estimation precision of various algorithms. Other conditions: DOA set $\{-5.5^\circ, 35.1^\circ, 60.2^\circ\}$ (ensure off grid). In Figure 3-4, four conclusions can be drawn (1) the RMSE performance of proposed method is most excellent; (2) proposed method depend less on number of snapshots than others; (3) the estimation performance of proposed method improves most with SNR increasing; (4) proposed method achieves perfect estimation performance.

5. Conclusions

In this paper, we further expand the use and performance of sparse Bayesian learning with respect to wideband signal by transforming MMV model into block-sparse model. The main contributions are as below: the hierarchical Bayesian prior framework is constructed to favour sparsity; the corresponding iterative process is derived and completed; the whole method is proposed and applied to solve underdetermined DOA estimation problems. In some way, our proposed method fully succeeds the advantages of sparse Bayesian learning and gains more excellent estimation performance

and sparse recovery ability. Simulation results of two experiments are presented to prove the superior of our method, such as favouring sparsity, joint sparsity, high estimation accuracy, stable estimated values and strong adaptability.

References

1. Krim H, Viberg M. Two decades of array signal processing research: the parametric approach. *IEEE SIGNAL PROC MAG*, **1996**, *13*, 67-94.
2. Liu C, Zakharov Y V, Chen T. Broadband underwater localization of multiple sources using basis pursuit de-noising. *IEEE T SIGNAL PROCES*, **2011**, *60*, 1708-1717.
3. Liu Z M, Huang Z T, Zhou Y Y. An efficient maximum likelihood method for direction-of-arrival estimation via sparse Bayesian learning. *IEEE T WIREL COMMUN*, **2012**, *11*, 1-11.
4. Su G, Morf M. The signal subspace approach for multiple wide-band emitter location. *IEEE T WIREL COMMUN*, **1983**, *31*, 1502-1522.
5. Allam M, Moghaddamjoo A. Two-dimensional DFT projection for wideband direction-of-arrival estimation. *IEEE T SIGNAL PROCES*, **1995**, *43*, 1728-1732.
6. Wang H, Kaveh M. Coherent signal-subspace processing for the detection and estimation of angles of arrival of multiple wide-band sources. *IEEE Trans. Acoust, Speech, Signal Processing*, **1985**, *33*, 823-831.
7. Hung H, Kaveh M. Focussing matrices for coherent signal-subspace processing. *IEEE Trans. Acoust, Speech, Signal Processing*, **1988**, *36*, 1272-1281..
8. El-Keyi A, Kirubarajan T. Adaptive beamspace focusing for direction of arrival estimation of wideband signals. *SIGNAL PROCESS*, **2008**, *88*, 2063-2077
9. Tang Z, Blacchiere G, Leus G. Aliasing-free wideband beamforming using sparse signal representation. *IEEE T SIGNAL PROCES*, **2011**, *59*, 3464-3469.
10. Shen Q, Liu W, Cui W, et al. Underdetermined DOA estimation under the compressive sensing framework: A review. *IEEE ACCESS*, **2016**, *4*, 8865-8878.
11. Gan L, Wang X. DOA estimation of wideband signals based on slice-sparse representation. *EURASIP J ADV SIG PR*, **2013**, *2013*, 1-10.
12. Qin Y, Liu Y, Liu J, et al. Underdetermined wideband DOA estimation for off-grid sources with coprime array using sparse Bayesian learning. *SENSORS-BASEL*, **2018**, *18*, 253.
13. Hyder M M, Mahata K. Direction-of-Arrival Estimation Using a Mixed $l_{2,0}$ Norm Approximation. *IEEE T SIGNAL PROCES*, **2010**, *58*, 4646-4655.
14. Malioutov D, Cetin M, Willsky A S. A sparse signal reconstruction perspective for source localization with sensor arrays. *IEEE T SIGNAL PROCES*, **2005**, *53*, 3010-3022.
15. He Z Q, Shi Z P, Huang L, et al. Underdetermined DOA estimation for wideband signals using robust sparse covariance fitting. *IEEE SIGNAL PROC LET*, **2014**, *22*, 435-439.
16. Hu N, Sun B, Zhang Y, et al. Underdetermined DOA estimation method for wideband signals using joint nonnegative sparse Bayesian learning. *IEEE SIGNAL PROC LET*, **2017**, *24*, 535-539.
17. Zhao L, Li X, Wang L, et al. Computationally efficient wide-band DOA estimation methods based on sparse Bayesian framework. *IEEE T VEH TECHNOL*, **2017**, *66*, 11108-11121.
18. Wipf D P, Rao B D. Sparse Bayesian learning for basis selection. *IEEE T SIGNAL PROCES*, **2004**, *52*, 2153-2164.
19. Wipf D P, Rao B D. An empirical Bayesian strategy for solving the simultaneous sparse approximation problem. *IEEE T SIGNAL PROCES*, **2007**, *55*, 3704-3716.
20. Zhang Z, Rao B D. Sparse signal recovery with temporally correlated source vectors using sparse Bayesian learning. *IEEE J-STSP*, **2011**, *5*, 912-926.
21. Tzikas D G, Likas A C, Galatsanos N P. The variational approximation for Bayesian inference. *IEEE SIGNAL PROC MAG*, **2008**, *25*, 131-146..